Probabilistic complex phase representation objective function for multimodal image registration

by

Alexander Wong

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2010

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

An interesting problem in computer vision is that of image registration, which plays an important role in many vision-based recognition and motion analysis applications. Of particular interest among data registration problems are multimodal image registration problems, where the image data sets are acquired using different imaging modalities. There are several important issues that make real-world multimodal registration a difficult problem to solve. First, images are often characterized by illumination and contrast non-uniformities. Such image non-uniformities result in local minima along the convergence plane that make it difficult for local optimization schemes to converge to the correct solution. Second, real-world images are often contaminated with signal noise, making the extraction of meaningful features for comparison purposes difficult to accomplish. Third, feature space differences make performing direct comparisons between the different data sets with a reasonable level of accuracy a challenging problem. Finally, solving the multimodal registration problem can be computationally expensive for large images.

This thesis presents a probabilistic complex phase representation (PCPR) objective function for registering images acquired using different imaging modalities. A probabilistic multi-scale approach is introduced to create image representations based on local phase relationships extracted using complex wavelets. An objective function is introduced for assessing the alignment between the images based on a Geman-McClure error distribution model between the probabilistic complex phase representations of the images. Experimental results show that the proposed PCPR objective function can provide improved registration accuracies when compared to existing objective functions.

Acknowledgements

First, let me start by thanking God for giving me strength in my research. Secondly, I wish to thank my supervisors Professor David Clausi and Professor Paul Fieguth for their advice and moral support throughout my time under their supervision. They have contributed tremendously to my growth as a dedicated researcher.

I wish to thank all of the funding agencies and corporations that supported my research through financial and knowledge support. In particular, I wish to thank the Natural Sciences and Engineering Research Council (NSERC) of Canada and the University of Waterloo Faculty of Engineering.

Finally, I wish to thank my family, my friends, my committee members, my colleagues at University of Waterloo, and CatMimi, Mimi, and Cici for their unconditional support.

Contents

List of Tables

List of Figures

List of Symbols

f	image being registered
g	reference image
$\underline{x}$	spatial coordinates
T	geometric transform
Ψ	objective function
s	wavelet scale
θ	wavelet orientation
J	wavelet function
J^o	odd-symmetric wavelet component
J^e	even-symmetric wavelet component
A	complex amplitude
ϕ	complex phase
w_o	center frequency
B	bandwidth
ϕ	mean phase
R	feature representation
E	energy
W	bilateral soft thresholding weighting function
∇	gradient
ϵ	divide-by-zero constant
ϱ	local neighborhood

w	scale space weighting function
L	scale representation
t	scaling parameter
κ	diffusion constant
c	conduction coefficient
$\sigma^2(\cdot)$	local spatial variance
$\lvert \cdot \rvert$	Euclidean spatial distance
p	true sampling distribution
q	instrumental sampling distribution
$\underline{\chi}$	drawn sample
τ	rejection threshold
ε	error residual

Chapter 1

Introduction

Image registration is the process of matching points in one image to their corresponding points in another image. The problem of image registration plays a very important role in many visual and object recognition and motion analysis applications. Some of these applications include visual motion estimation [1, 2], vision-based content-based retrieval [3, 4], image registration [5–8], and biometric authentication [9]. In the best case scenario, the images exist at the same scale, in the same orientation, as well as represented in the same feature space. However, this is not the case in most real-world applications. There are many situations where the images exist in different feature spaces. This particular problem will be referred to as the multimodal registration problem and is a particularly difficult problem to solve. Examples of this problem in real-world situations include medical image registration and tracking of MRI/CT/PET data [10] and building modeling and visualization using LIDAR and optical data [11, 12]. In the past, the task of multimodal registration was performed manually by placing markers on the scene during the image acquisition process and then finding the geometric transformation that brings the markers into alignment. This approach to multimodal registration is very laborious and time consuming. As such, methods for registering images acquired using different modalities in an automated fashion is highly desired.

While current methods differ in their specific approach to the automatic image registration problem, the majority of these methods can be broken down into the following steps:

1. **Feature representation**: Global and/or local descriptors are extracted from the set of images under evaluation based on a particular feature representation.

2. **Similarity analysis**: Similarity analysis is performed between the set of image descriptors using an objective function.

3. **Transformation estimation**: Based on the similarity analysis, the transformation model is estimated to provide the best alignment between the images.

4. **Transformation and resampling**: The images are transformed based on the determined model and are resampled using an interpolation method.

Each of these steps plays an important part in the image registration procedure. Therefore, the following criteria are important in the design of an automatic multimodal image registration framework:

1. **Efficiency**: Certain applications such as medical and remote sensing imaging analysis require the registration of large images. This is further complicated when registering large volumes of image data. Therefore, to minimize the computational effort required to perform each of these steps while maintaining alignment accuracy is important.

2. **Robustness**: Differences in multimodal images of the same scene exist due to factors such as environmental noise, differences in illumination and contrast, differences in viewpoint and differences in imaging modalities. Therefore, to minimize the effect of such image variances on image registration accuracy is important.

3. **Accuracy**: Visualization and analysis of images typically require that a reasonable level of accuracy be achieved during the registration process. Therefore, the registration process should produce an image that is visually and numerically correct.

The goal of the thesis is to design and develop an objective function that satisfies the criteria associated with robustness and accuracy for the purpose of registering images acquired using different imaging modalities.

1.1 Multimodal Registration Problem

The multimodal registration problem can be defined in the following manner. Suppose there exist two images f and g, each containing N_f and N_g points respectively. Furthermore, points in f and g are represented using two different feature spaces. For every point

in f, the goal of registration is to determine a corresponding point in g such that f and g are aligned. This problem can alternatively be formulated as finding the optimal transformation T that maps all points from f to the points from g such that the alignment between f and g is achieved, resulting in the following optimization problem:

$$T_{\text{opt}} = \text{argmax}_T \left[\Psi \left(g(\underline{x}), f(T(\underline{x}))\right)\right] \tag{1.1}$$

where T_{opt} is the optimal transform and $\Psi(.)$ is the objective function assessing similarity.

1.2 Challenges of Multimodal Registration

There are many important issues that make automated multimodal registration a very challenging problem to solve.

1. **Geometric Distortions**: First, images being acquired using different modalities (i.e., MR and CT) are captured using different imaging devices at different times. As such, the images acquired using different modalities often contain different geometric distortions that make it difficult to compare image content.

2. **Noise**: Real-world images are also characterized by point noise, making the extraction of meaningful features for comparison purposes difficult to accomplish.

3. **Intensity Mapping Differences**: Images acquired using different modalities are mapped to different intensity values. This makes it difficult to compare images based on their intensity values since the same content within the images may be represented by different intensity values. The problem is further complicated by the fact that various intrinsic (for MRI, static field and RF non-homogeneity [13, 14]) and extrinsic (patient motion) sensing conditions can lead to image non-uniformities. Such image non-uniformities result in the same content within a single image to be represented by different intensity values.

4. **Convergence Issues**: The disparity between the intensity values of multimodal images can lead to coincidental local intensity matches between non-corresponding content. These matches can result in local minima that make it difficult for local optimization schemes such as conjugate gradient [15], Nelder-Mead simplex [16], and sequential quadratic programming (SQP) methods [17] to converge to the correct

solution, particularly when the initial overlap is small and the distance traveled during the optimization process is large.

5. **Computational Complexity**: Solving the multimodal registration problem can be computationally expensive for large images.

The goal of the proposed objective function is to address the issues associated with noise, intensity mapping differences, as well as convergence.

1.3 Thesis Contributions

The thesis makes the following contributions:

- A novel approach to structure-based image representations based on local phase relationships extracted using complex wavelets was introduced, which is described in detail in Chapter 3,
- a novel probabilistic multi-scale approach is introduced for constructing structure-preserving and noise-resilient multi-scale representations of images, which is described in detail in Chapter 4, and
- a robust objective function is introduced for determining the alignment between the images based on a Geman-McClure error distribution model between the probabilistic complex phase representations of the images, which is described in detail in Chapter 5.

Chapter 2

Overview of Automatic Registration Methods

This chapter provides an overview of existing research literature related to automatic image registration, as well as a discussion of the advantages and disadvantages of viable solutions for multimodal image registration. The background literature pertaining to the individual concepts behind the proposed objective function for multimodal image registration is provided in subsequent chapters. An overview of existing complex wavelet representations is provided in Section 3.2. An overview of existing scale space representations is provided in Section 4.2. Finally, an overview of existing estimation functions used for estimating the alignment between images is provided in Section 5.3.

2.1 Types of Registration Methods

A large number of methods have been proposed for the purpose of image registration, where the underlying goal is to find the optimal alignment between a set of images. In general, current methods can be grouped into four main types:

1. Methods based on relative distances [18–23]
2. Methods in the frequency domain [24–26]
3. Methods based on direct comparisons [5, 6, 8, 27–39]

4. Methods based on extracted features [40–50]

2.2 Relative Distance Methods

Methods based on relative distances exploit the spatial relationships between neighboring pixels within an image to determine the best match between two points. These algorithms are based on the assumption that if a point $p_{f,0}$ in image f has a corresponding point $p_{g,0}$ in image g, then there exist other points $\{p_{f,1}, p_{f,2}, ..., p_{f,n}\}$ in f that have corresponding points $\{p_{g,1}, v_{g,2}, ..., p_{g,n}\}$ in g, such that the distance between $p_{f,0}$ and point $p_{f,k}$ is equal to the distance between $p_{g,0}$ and point $p_{g,k}$. Methods based on relative distances can be further subdivided into two types: i) relaxation-based methods, and ii) iterated closest point (ICP) methods. In relaxation-based methods [18–20], a support value is calculated between each possible point pair and a match is made if it yields the best support from other points. In ICP-based methods [21–23], points from f and g are first matched to each other using a nearest neighbor algorithm. Parameter estimation is then performed based on these initial point pairs to determine a transformation that maps points from f to g. The points in f are then transformed using the estimated parameters and the entire process is repeated until a residual error condition is satisfied. As such, ICP-based methods require a good initial transformation estimation to produce reasonable results. Methods based on relative distances are primarily useful for situations where the transformation between the images consists only of translations and rotations.

2.3 Frequency Domain Methods

Methods in the frequency domain [24–26] exploit the frequency characteristics such as phase to estimate the transformation between two images. A common frequency domain registration method is phase correlation, where the Fourier coefficients calculated from image f are divided by those calculated from image g. Performing the inverse transform on the result yields a single peak indicating the translation that matches the two images. This technique has been extended to account for global rotations and scale by Reddy et al. [25]. As such, frequency domain methods are currently only suited for globally rigid point correspondences (e.g., scale, rotation, and translation).

2.4 Direct Comparison Methods

Methods based on direct comparisons [5, 6, 8, 27–39] attempt to find point correspondences between two images by performing matches directly in their respective feature spaces. Many techniques in this group make use of feature information from neighboring points to determine the similarity between two points. Some common similarity metrics used in direct comparison methods include maximum likelihood [7], correlation [28, 31], correlation ratio [33–35], and mutual information [5, 6, 27, 8, 38, 39].

Of particular interest in recent years are techniques based on mutual information, which attempt to match points by finding the mutual dependence between the two images. An objective function based on mutual information can be expressed as follows:

$$\Psi\Big(g\left(\underline{x}\right), f\left(T\left(\underline{x}\right)\right)\Big) = H\Big(g\left(\underline{x}\right)\Big) + H\Big(f\left(T\left(\underline{x}\right)\right)\Big) - H\Big(g\left(\underline{x}\right), f\left(T\left(\underline{x}\right)\right)\Big) \tag{2.1}$$

where $H(\cdot)$ denotes the marginal intensity entropy and $H(\cdot,\cdot)$ denotes the joint intensity entropy. In this case, maximizing the objective function maximizes mutual information, which in effect minimizes the joint intensity entropy. Methods that utilize mutual information based objective functions take advantage of the fact that correctly registered images are characterized by tightly packed joint distributions and that minimizing the joint intensity entropy effectively minimizes the dispersion of the joint distribution. The main advantage of mutual information is that it allows for the direct intensity-based comparison of multimodal images, making no underlying assumptions regarding the intensity relationships between the images under evaluation.

2.5 Feature-based Methods

In methods based on extracted features [40–50], the images are transformed into a common feature space before they are compared. This is in contrast to direct comparison methods, which match points in their respective feature spaces. Methods based on extracted features attempt to find point matches between two images by indirectly comparing points using extracted features that exist in a common feature space. Common features used in such methods include edges [50], contours [43, 44], invariant moments [46], orientation [48], and shape properties [42]. By comparing methods using derived features in a common feature space, this allows such techniques to match data images existing in different feature spaces in cases where similar features can be extracted from both images. A particularly interesting case is the method proposed by Mellor et al. [40, 41], which extracts the local

phase from the images and finds point matches between them using mutual information. Local phase extracted from multimodal images in itself cannot be compared in a direct manner. However, since the local phase is highly invariant to feature non-uniformities that may occur in real-world situations (such as illumination and contrast non-uniformity in the case of imaging data) and greatly emphasizes the structural characteristics of the underlying data, the use of local phase can greatly improve the performance of mutual information. This method can be considered one of the state-of-the-art techniques for matching multimodal data.

2.6 Discussion

Based on the above descriptions, the methods best suited for registration of multimodal image data appear to be those based on extracted features and those based on mutual information. However, there are a number of issues associated with these registration methods. The major drawback associated with methods based on extracted features is that they require the extraction of features within a common feature space. While this allows such methods to match images existing in different feature spaces, it requires similar features to be extracted from both images for them to work effectively.

There are also several drawbacks to the use of mutual information based methods. First, since the intensity relationship between the images is relatively unconstrained by mutual information, the convergence planes associated with mutual information based methods possess high non-monotonicity with many local optima [34]. Some causes that can lead to high non-monotonicity include sampling (i.e., number of histogram bins and image resolution) and interpolation effects. Since local optimization schemes are dependent on the monotonicity of the underlying cost function, such schemes can often get trapped in local maxima. This is particularly problematic in situations with small initial overlaps, where the optimization scheme must travel a long distance to the correct solution. Direct methods based on mutual information require good initial estimates for good results. However, this is not always possible, particularly in cases where a large volume of images need to be matched. This need for careful initialization of the system to achieve proper registration is one of the drawbacks of using mutual information. Furthermore, techniques based on mutual information are computationally expensive as they require the calculation of marginal and joint entropies. As such, mutual information based methods may not be practical for certain situations where computational speed is important. More importantly, mutual information based objective functions suffer from what is referred to as the "curse of di-

mensionality" [51]. As the number of dimensions increase, the number of samples required to reliably approximate entropy grows exponentially. Therefore, this makes it very difficult to extend entropy-based methods to high-dimensional multimodal registration problems as there are typically insufficient number of samples to estimate entropy properly.

In addition to the issues associated with these registration methods, it is important to deal with factors such as environmental noise and contrast non-uniformities due to imaging modalities. The proposed objective function aims to address the issues associated with methods based on extracted features, as well as noise and contrast non-uniformities.

Chapter 3

Complex Phase Representation

This chapter describes in detail the overall design and performance analysis of the proposed complex wavelet representation. The computation of complex amplitude and phase information using complex wavelet transforms is presented. An overview of existing complex wavelet representations is presented to examine the challenges and issues that need to be addressed in the design of complex wavelet representations. An image representation based on complex phase relationships is described in detail. The structural sensitivity of the complex phase representation is studied and novel methods for improving structural sensitivity is proposed. Finally, the noise sensitivity of the complex phase representation is studied and novel methods for reducing sensitivity to signal noise is proposed. The main contributions in this chapter are: i) the introduction of a new phase order response weighting function for improving the response sensitivity to structural characteristics for improved structural contrast, as described in Section 3.4, and ii) the introduction of a new bilateral soft threshold scheme for reducing noise sensitivity, as described in Section 3.5.

3.1 Complex Amplitude and Phase

To compute the complex amplitude and phase information of an image f, multi-scale complex wavelet representations of f are constructed using a complex wavelet transform such as the Gabor complex wavelet transform and the Log-Gabor complex wavelet transform [52, 53], with each of the α scales (s) and β orientations (θ) maintaining the resolution of the original images.

One of the most popular complex wavelet transforms used for feature extraction is

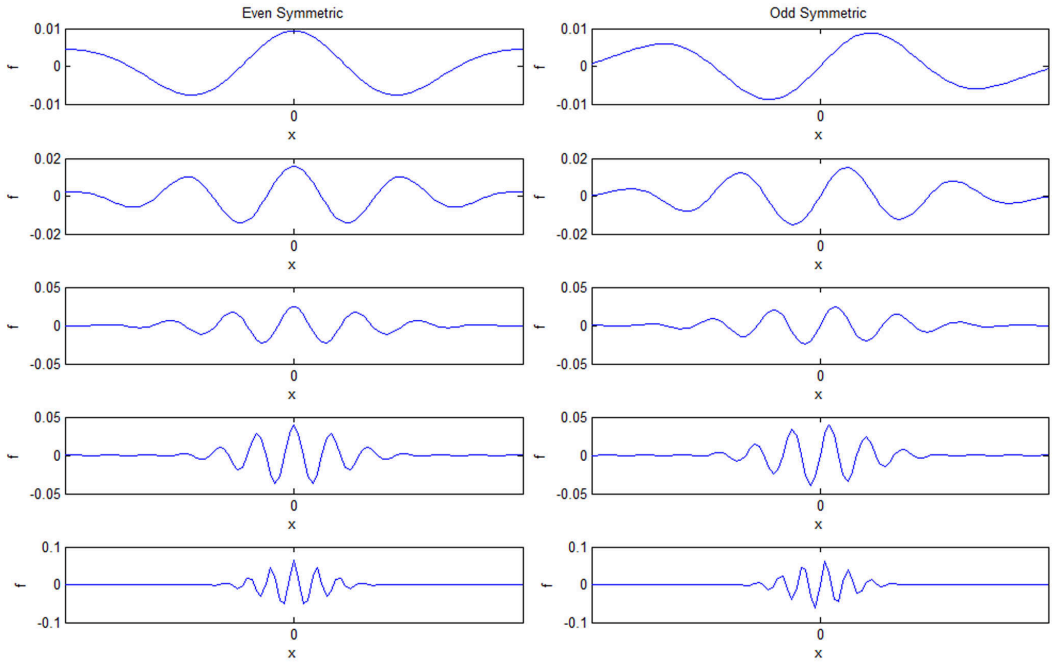

Figure 3.1: Gabor complex wavelets (Eq. (3.1)) at increasing frequencies (from top to bottom) with an octave bandwidth. The even- and odd-symmetric quadrature pair are shown for each complex wavelet.

the Gabor complex wavelet transform. The Gabor wavelet can be defined as the product between a harmonic function and a Gaussian function,

$$J(x) = \left(\frac{1}{2\pi\sigma} \mathrm{e}^{\frac{-x^2}{2\sigma^2}} \right) e^{j2\pi w_o x}, \tag{3.1}$$

where w_o is the center frequency and σ controls the standard deviation of the Gaussian envelope. Five Gabor complex wavelets at different frequencies with an octave bandwidth are shown in Fig. 3.1. One of the benefits of the Gabor complex wavelet transform is that it provides excellent localization of spatial-frequency information at different orientations.

A limitation of Gabor complex wavelets described in Eq. (3.1) is that the zero mean criterion cannot be achieved for bandwidths over one octave [54]. This limitation of Gabor complex wavelets is illustrated in Fig. 3.2, which shows the even-symmetric components of three Gabor complex wavelets with bandwidths of two octaves and their corresponding amplitude spectra. Each of the even-symmetric components have non-zero DC components,

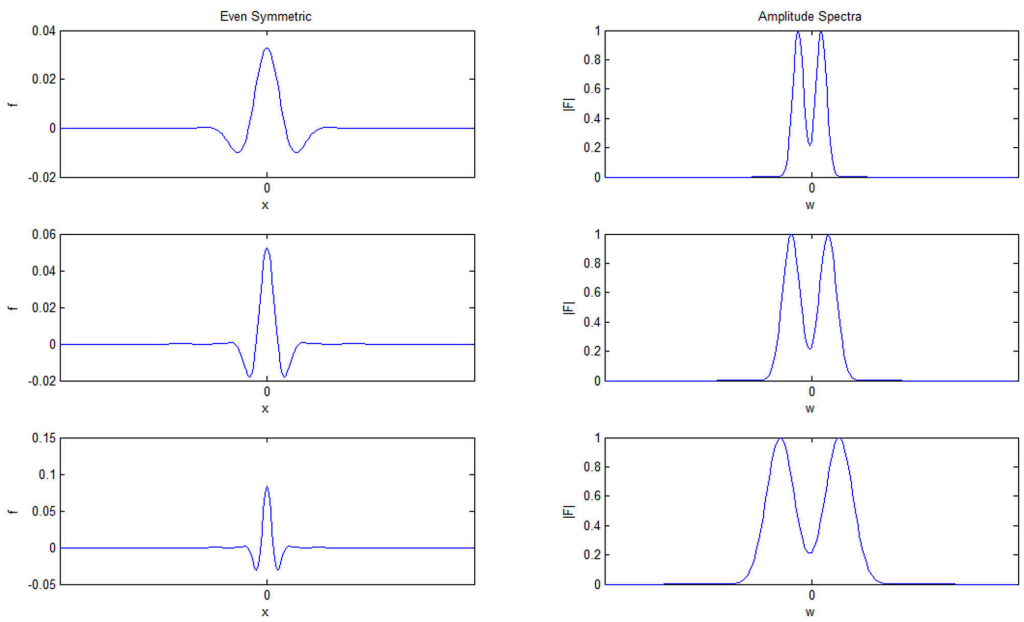

Figure 3.2: The even-symmetric components of three Gabor complex wavelets with bandwidths of two octaves and their corresponding amplitude spectra. The amplitude spectra of each even-symmetric component have non-zero DC components, thus violating the fundamental zero mean criterion of wavelets.

thus violating the fundamental zero mean criterion of wavelets. Therefore, Gabor complex wavelets are restricted to bandwidths below one octave and hence a large number of Gabor complex wavelets are necessary for wide spectral coverage. This makes Gabor complex wavelets poorly suited for efficient representation of broad spectral information.

To achieve efficient representation of broad spectral information while attaining excellent spatial-frequency information like the Gabor complex wavelet transform, one effective approach is to utilize the Log-Gabor complex wavelet transform proposed by Field [52]. Based on a statistical analysis of natural imagery, Field observed that the amplitude spectra of such imagery is highest at low frequencies and decays rapidly by a factor of approximately $\frac{1}{w}$. Intuitively, a complex wavelet with similar amplitude spectra characteristics as the imagery would be better suited to represent such imagery in an efficient manner. The Fourier transform of the Log-Gabor complex wavelet can be expressed as

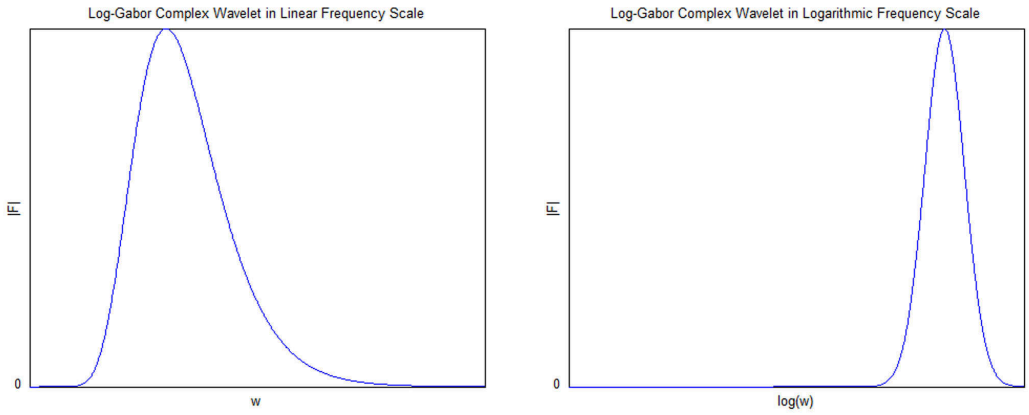

Figure 3.3: An example amplitude spectrum of a Log-Gabor complex wavelet (Eq. (3.2)) in both linear and logarithmic frequency scales.

$$J(w) = \exp\left[\frac{-\log\left(w/w_o\right)^2}{2\log\left(B\right)^2}\right], \tag{3.2}$$

where B determines the bandwidth of the complex wavelet. An example amplitude spectrum of a Log-Gabor complex wavelet in both linear and logarithmic frequency scales is shown in Fig. 3.3. There are several important observations that can be made about the Log-Gabor complex wavelet. First, the DC component of the Fourier transform of a Log-Gabor complex wavelet is always zero. This guarantees that any Log-Gabor complex wavelet will satisfy the zero mean criterion of wavelets. Therefore, the Log-Gabor complex wavelet transform allows for the construction of wavelets with arbitrarily large bandwidths to provide wide spectrum coverage with fewer wavelets than the Gabor complex wavelet transform is capable of. Second, the amplitude spectra of the Log-Gabor complex wavelet decays rapidly at the high frequencies, which is similar to that observed by Field for natural imagery. Hence, Log-Gabor complex wavelets can potentially provide more efficient representation for such imagery. Finally, the Log-Gabor complex wavelet has a symmetric Gaussian shape in the logarithmic frequency scale. This symmetry in the logarithmic frequency scale is consistent with the response properties of mammalian cortical cells [52].

Fig. 3.4 shows the Log-Gabor complex wavelets at the same frequencies and bandwidth as the Gabor complex wavelets shown in Fig. 3.1. The Log-Gabor complex wavelets are very similar to the Gabor complex wavelets. As a contrast to the amplitude spectra

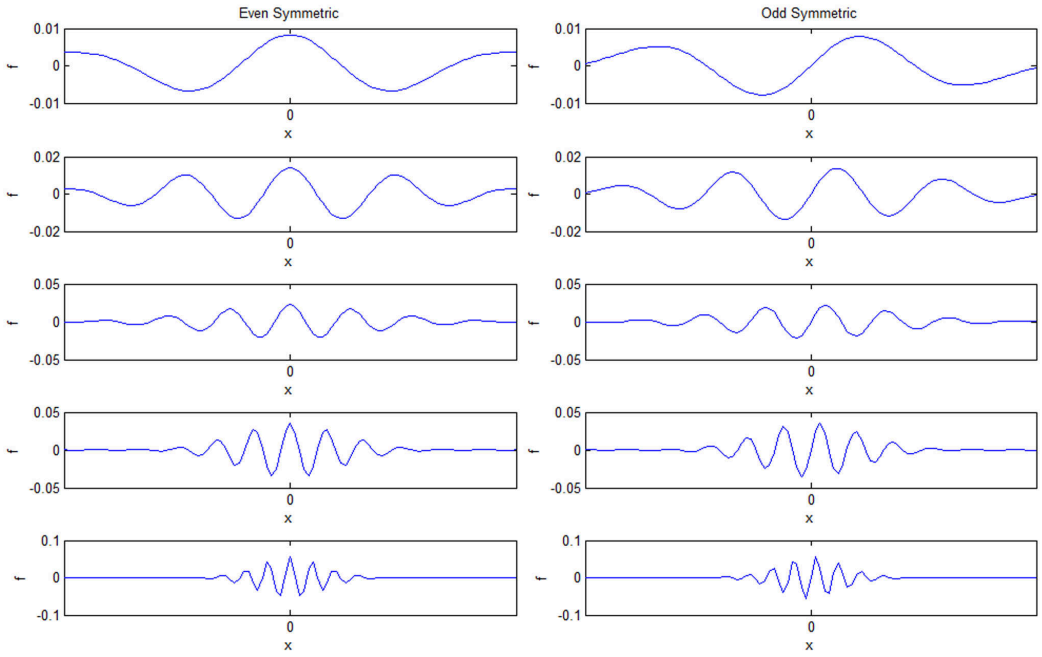

Figure 3.4: Log-Gabor complex wavelets at the same frequencies and bandwidth as the Gabor complex wavelets shown in Fig. 3.1. The even- and odd-symmetric quadrature pair are shown for each complex wavelet.

shown in Fig. 3.2, three even-symmetric components of Log-Gabor complex wavelets with bandwidths of two octaves and their corresponding amplitude spectra is shown in 3.5. The overall shapes of the Log-Gabor complex wavelets are similar to that of the Gabor complex wavelets shown Fig. 3.2, but with sharper peaks and extended tails. The sharper peaks and extended tails of the Log-Gabor complex wavelet allows the zero mean criterion to be satisfied, as seen in the DC component of the corresponding amplitude spectra. As such, the use of the Log-Gabor complex wavelet transform allows for wider spectral coverage while still providing high spatial localization like Gabor complex wavelets.

For the proposed framework, the Log-Gabor complex wavelet transform was performed over $\alpha = 4$ scales and $\beta = 6$ orientations, with wavelengths of 3, 9, 27, and 81 pixels to maintain bandwidths of two octaves. Each point $\underline{x}$ in the multi-scale complex wavelet representation is represented by a set of $\alpha\beta$ complex wavelet responses $\Upsilon_{s,\theta}$, where

$$\Upsilon_{s,\theta} = A_{s,\theta}(\underline{x})e^{j\phi_{s,\theta}(\underline{x})}. \tag{3.3}$$

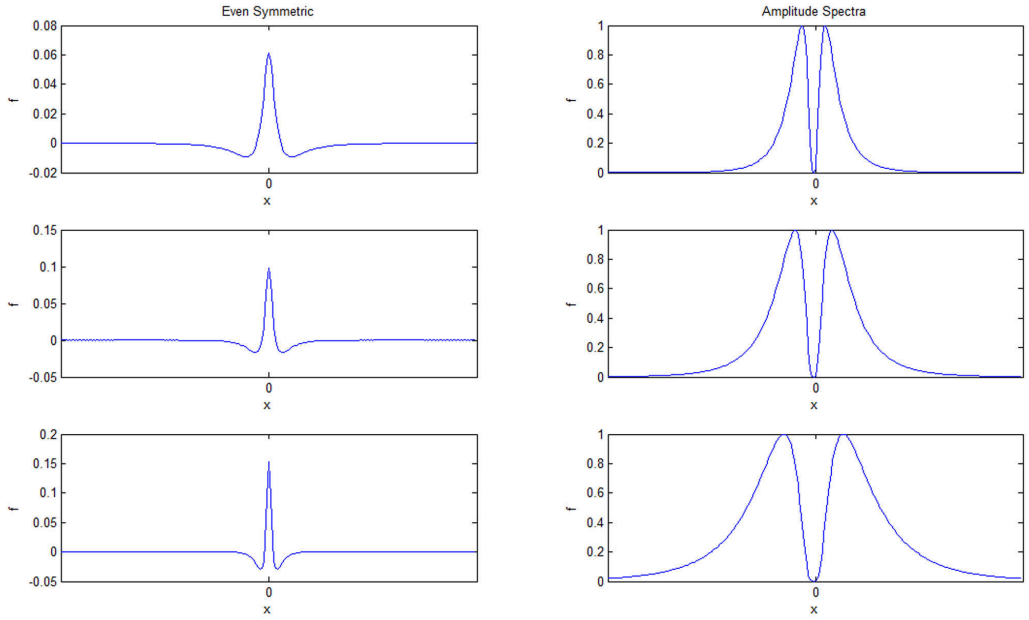

Figure 3.5: The even-symmetric components of three Log-Gabor complex wavelets with bandwidths of two octaves and their corresponding amplitude spectra. The overall shapes of the Log-Gabor complex wavelets are similar to that of the Gabor complex wavelets shown Fig. 3.2, but with sharper peaks and extended tails. The sharper peaks and extended tails of the Log-Gabor complex wavelet allows the zero mean criterion to be satisfied, as seen in the DC component of the corresponding amplitude spectra.

The amplitude $A_{s,\theta}(\underline{x})$ and phase $\phi_{s,\theta}(\underline{x})$ for a complex wavelet response $\Upsilon_{s,\theta}$ presented in Eq. (3.3) can be computed as

$$A_{s,\theta}(\underline{x}) = \sqrt{\left(f(\underline{x}) * J^e_{s,\theta}(\underline{x})\right)^2 + \left(f(\underline{x}) * J^o_{s,\theta}(\underline{x})\right)^2}, \tag{3.4}$$

and,

$$\phi_{s,\theta}(\underline{x}) = \tan^{-1}\left(\frac{f(\underline{x}) * J^e_{s,\theta}(\underline{x})}{f(\underline{x}) * J^o_{s,\theta}(\underline{x})}\right), \tag{3.5}$$

where $J^e_{s,\theta}(\underline{x})$ and $J^o_{s,\theta}(\underline{x})$ are the even- and odd-symmetric Log-Gabor quadrature pairs, and $*$ denotes a convolution.

3.2 Complex Wavelet Representations

Images acquired using different modalities have different intensity mappings for the same content, making them difficult to compare in a direct manner. One possible solution to this problem is to construct representations of the multimodal images that allow for comparison independent of the underlying image intensity. A common approach to an intensity-independent representation of the images is based on its structural characteristics such as edges and corners. This type of approach assumes that two images obtained of the same scene using different modalities can have significantly different intensity characteristics but should have very similar structural characteristics. Therefore, based on this assumption, the structural characteristics of the disparate images can be compared in a direct fashion. While previous attempts have been made to develop structural representations using gradient fields [48, 50], these methods are sensitive to image non-uniformity and can lead to incorrect alignments. Therefore, such structural feature representations are not suitable for situations characterized by image non-uniformities.

In recent years, considerable interest has been shown towards developing structural feature representations based on complex wavelets [40, 41, 54–60]. This is partly motivated by the fact that complex wavelets have been shown to provide a good basic model of the human visual cortex, which is effective at identifying structural significance in visual information under various conditions [55, 61–63]. This makes such representations very well suited for the purpose of multimodal image registration, where the same scene is represented using different intensity mappings but have the same underlying structure. Another motivation for developing feature representations based on complex wavelets is that, unlike feature extraction methods such as those proposed in [42, 44, 45, 48, 50], they allow for structural analysis across multiple scales, thus allowing for extraction of both fine and coarse structural details from the visual information. Given these benefits, several feature representations have been proposed based on complex wavelets. One of the earliest complex wavelet feature representations is based on the Fourier representation [55], which can be defined as the normalized weighted summation of cosine-weighted complex phase deviations from the amplitude-weighted mean phase $\bar{\phi}\left(\underline{x}\right)$ across α scales,

$$R(\underline{x}) = \max_{\bar{\phi}\in[0,2\phi]} \frac{\sum_{s=1}^{\alpha} A_s\left(\underline{x}\right)\cos\left(\phi_s\left(\underline{x}\right) - \bar{\phi}\left(\underline{x}\right)\right)}{\sum_{s=1}^{\alpha} A_s\left(\underline{x}\right) + \epsilon}. \tag{3.6}$$

where s denotes the wavelet scale and ϵ is a small constant used to avoid division by

zero. An alternative formulation of this feature representation was proposed by Venkatesh and Owens [64], which is defined as the the ratio between the energy of complex wavelet components E and the sum of amplitudes A_s,

$$R(\underline{x}) = \frac{E(\underline{x})}{\sum_{s=1}^{\alpha} A_s(\underline{x}) + \epsilon}, \tag{3.7}$$

where s denotes the wavelet scale and ϵ is a small constant used to avoid division by zero. While this representation is robust to illumination and contrast non-uniformities, this feature representation is very sensitive to noise and provides poor localization of structural information [65]. To account for the poor structural localization issues associated with the representation proposed by Morrone et al., Kovesi proposed the use of local phase order at different orientations θ for the purpose of image representation [65]. The basic image representation proposed by Kovesi can be expressed as the normalized summation of phase deviations across different scales and orientations,

$$R(\underline{x}) = \frac{\sum_{q=1}^{\beta} \sum_{s=1}^{\alpha} W_{\theta_q}(\underline{x}) \left\lfloor A_{s,\theta_q}(\underline{x}) \, \Delta\Phi_{\theta_q}(\underline{x}) - \tau \right\rfloor}{\sum_{q=1}^{\beta} \sum_{s=1}^{\alpha} A_{s,\theta_q}(\underline{x}) + e} \tag{3.8}$$

$$\Delta\Phi_{\theta}(\underline{x}) = \cos\left(\phi_{s,\theta}(\underline{x}) - \bar{\phi}_{\theta}(\underline{x})\right) - \left|\sin\left(\phi_{s,\theta}(\underline{x}) - \bar{\phi}_{\theta}(\underline{x})\right)\right|, \tag{3.9}$$

where $A_{s,\theta}$ and $\phi_{s,\theta}$ represent the amplitude and phase at wavelet scale s and orientation θ defined in Eq. (3.4) and Eq. (3.5) respectively, $\bar{\phi}_{\theta}$ represents the weighted mean phase at orientation θ, $\lfloor . \rfloor$ denotes a truncation operator that sets all negative values to zero, τ is a hard noise threshold, and e is a small constant used to avoid division by zero. This feature representation is robust to low levels of image noise and provides improved localization of structural information [57]. There are several drawbacks to this representation. First, the weighting function used has limited sensitivity to structural characteristics with high structural significance, which is important in capturing and distinguishing fine structural detail. The weighting function used by Kovesi also has some counterintuitive characteristics that affect the reliability of the feature representation. Furthermore, the method proposed by Kovesi for estimating local phase order runs into problems in situations characterized by low signal-to-noise ratios.

Mellor et al. [40, 41] proposed the use of local phase obtained using complex wavelets directly as a basis for feature representation.

$$\phi_n(\underline{x}) = \tan^{-1}\left(\frac{(f(\underline{x}) * J_n^e(\underline{x}))}{(f(\underline{x}) * J_n^o(\underline{x}))}\right) \tag{3.10}$$

where $J_n^e(\underline{x})$ and $J_n^o(\underline{x})$ are the even-symmetric and odd-symmetric wavelets at scale n. The key advantage to the use of local phase representations over other representations is the fact that it is largely independent of intensity. Thus, it is highly robust to the presence of image non-uniformities. There are major drawbacks to this representation. First, the local phase is not directly comparable for images acquired using different imaging modalities [66], thereby requiring more computationally complex techniques such as mutual information for similarity analysis. Secondly, these local phase representations provide no information on structural significance, thus placing no importance on significant structural characteristics. Finally, image noise is not accounted for.

Fauquer et al. [56] proposed a feature representation created based on the sum of products of subband amplitudes ρ extracted using Dual-tree complex wavelets [58]. While this representation accounts for image noise, it is solely based on amplitudinal information, making it sensitive to image non-uniformities. Liu et al. [59] and Hemmendorff et al. [60] proposed the use of local phase gradients for feature representation. The main disadvantage of such representations is that the computation of local phase gradients is heavily dependent on maximum local amplitude, making it sensitive to image non-uniformity. Furthermore, these representations do not account for image noise.

Based on the aforementioned complex wavelet representations, it can be seen that the main advantage to such methods is that they are highly robust to image non-uniformities, which is very important for the purpose of registering images acquired using different modalities. However, current representations are highly sensitive to situations characterized by low signal-to-noise ratios and some are sensitive to intensity mappings. Therefore, a complex wavelet representation that is highly robust to image noise is desired.

3.3 Representation Based on Complex Phase Relationships

Images acquired using different modalities have different intensity mappings for the same content, making them difficult to compare in a direct manner. One possible solution to this

problem is to construct representations of the multimodal images that allow for comparison independent of the underlying image intensity. A common approach to an intensity-independent representation of the images is based on its structural characteristics such as edges and corners. This type of approach assumes that two images obtained of the same scene using different modalities can have significantly different intensity characteristics but should have very similar structural characteristics. Therefore, based on this assumption, the structural characteristics of the disparate images can be compared in a direct fashion. While previous attempts have been made to develop structural representations using gradient fields, these methods are sensitive to image non-uniformity and can lead to incorrect alignments. Therefore, a structural representation that is relatively independent of image non-uniformity is desired.

The approach used in the proposed framework for constructing a structural representation of the images is to use local phase relationships extracted using complex wavelets to identify structurally significant characteristics within an image. Local phase relationships was recently proposed as an effective method for determining structural significance in an image [57, 63, 67]. This approach to measuring structural significance is based on the theory that the local phase order peaks at locations of high perceptual significance [63]. More importantly, these points of high perceptual significance coincide with points of high structural significance within an image. Furthermore, representations based on local complex phase relationships is largely independent of intensity. These properties make the use of local complex phase relationships an effective method for creating a structural representation of the images that can be evaluated in a direct fashion.

To capture the structural characteristics of images using local complex phase relationships, the representation proposed by Morrone et al. [55] is extended for use with the Log-Gabor complex wavelet by integrating across different complex wavelet orientations. For the proposed framework, the structural significance at a point $\underline{x}$ can be quantified as the normalized weighted summation of cosine-weighted complex phase deviations from the amplitude-weighted mean phase $\bar{\phi}\left(\underline{x}, \theta\right)$ across α scales and β orientations,

$$R(\underline{x}) = \frac{\sum_{q=1}^{\beta} \sum_{s=1}^{\alpha} A_{s,\theta_q}\left(\underline{x}\right) \cos\left(\phi_{s,\theta_q}\left(\underline{x}\right) - \bar{\phi}_{\theta_q}\left(\underline{x}\right)\right)}{\sum_{q=1}^{\beta} \sum_{s=1}^{\alpha} A_{s,\theta_q}\left(\underline{x}\right)}. \tag{3.11}$$

Based on Eq. (3.11), phase order across scales and orientations increases, corresponding to increasing structural significance, the cosine-weighted complex phase deviations approach

one and $R(\underline{x})$ approaches one as well. Conversely, as phase order decreases, corresponding to decreasing structural significance, the cosine-weighted complex phase deviations approach negative one and $R(\underline{x})$ approaches negative one as well. As such, it can be seen that complex phase relationships can be very effective at providing a image representation that is robust to intensity differences and contrast non-uniformities.

A simplified visualization for the concept of local complex phase order is illustrated in Fig. 3.6. Suppose that the local complex wavelet components (from different scales) at a particular point are plotted on a polar plot in the form of a vector sum. When the complex wavelet components are maximally in phase, the magnitude of the vector sum is equal to the sum of the individual amplitudes and the local phase order goes to one. Conversely, as the complex wavelet components go out of phase, the local phase order decreases.

To demonstrate the correspondence between local phase order and structural significance, as well as robustness to intensity differences and contrast non-uniformities, a multi-scale complex wavelet representations of a piece-wise step function are constructed using a Log-Gabor complex wavelet transform with four scales. The complex wavelet responses to the piece-wise step function (composed of two steps of different amplitudes) at different frequencies are shown in Fig. 3.7. The complex wavelet responses across the different frequencies exhibit maximum phase order at the edges of the piece-wise step function. Based on the complex wavelet responses, the complex wavelet phase representation of the piece-wise step function is computed and shown in Fig. 3.8. The representation peaks at the edges of the piece-wise step function where maximum phase order is achieved. This demonstrates that the correspondence between local phase order and structural significance holds true for simple structures such as step edges. It can also be observed that the representation peaks are similar in value, despite the fact that the steps they correspond to have significantly different amplitudes. This demonstrates that the complex phase representation is robust to intensity differences and contrast variations.

For a more complex demonstration of correspondence between local phase order and structural significance, as well as robustness to intensity differences and contrast non-uniformities, the complex wavelet phase representation of the well-known Shepp-Logan phantom test image is computed using a 2D Log-Gabor complex wavelet transform with four scales and six wavelet orientations and bandwidths of two octaves and shown in Fig. 3.9. The structural information from the phantom test image are well preserved in the complex phase representation. Furthermore, the structural significance indicated by the complex phase representation is robust to contrast and intensity differences in the image.

The key benefit of this complex phase representation is that such representations can

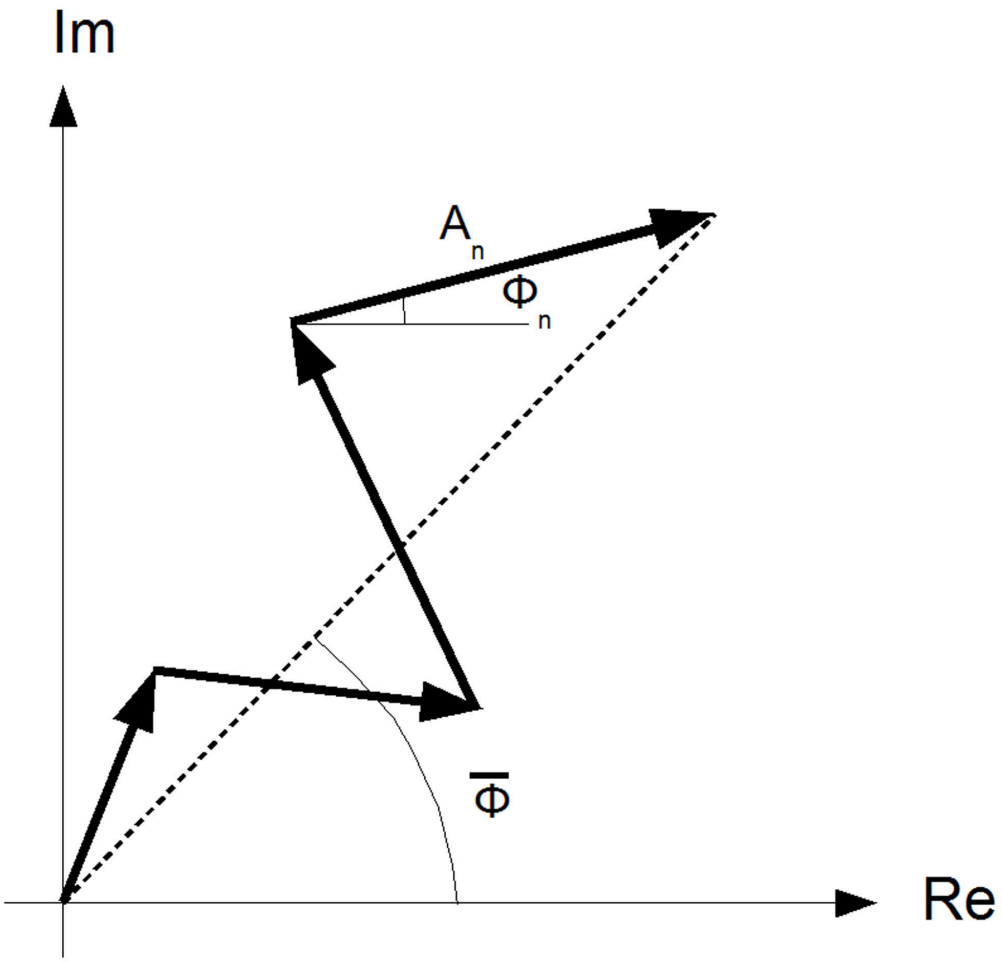

Figure 3.6: Simple visualization of local complex phase order. A_n and ϕ_n represents the amplitude and phase of the n^{th} wavelet component respectively. $\bar{\phi}$ represents the amplitude-weighted mean phase. When the complex wavelet components are maximally in phase, the magnitude of the vector sum (i.e., the dashed line) is equal to the sum of the individual amplitudes and the local phase order goes to one. Conversely, as the complex wavelet components go out of phase, the local phase order decreases.

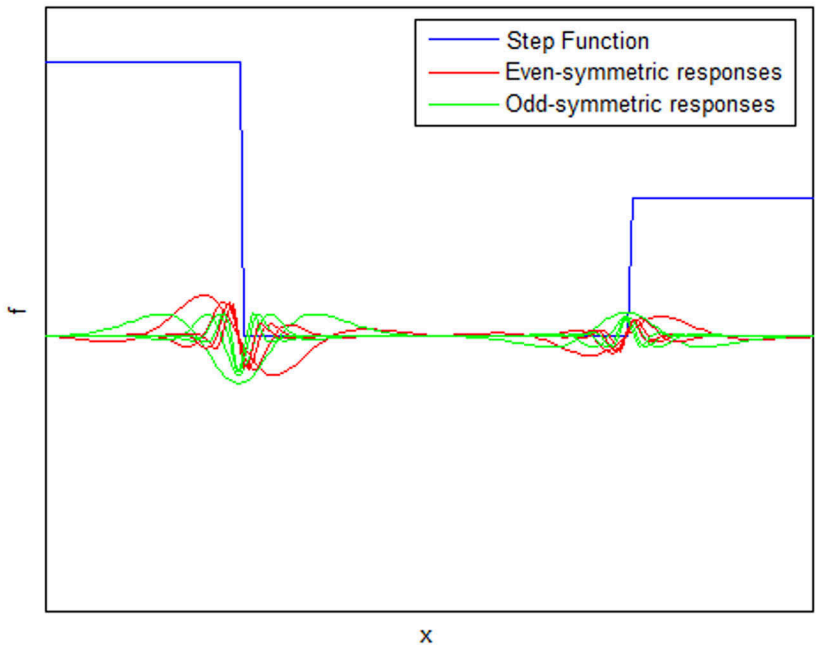

Figure 3.7: The complex wavelet responses (even-symmetric responses $f(x) * J^e_{s,\theta}(x)$ and odd-symmetric responses $f(x) * J^o_{s,\theta}(x)$) to the piece-wise step function $f(x)$ (composed of two steps of different amplitudes) at different frequencies.

be compared in a direct manner while intensity representations rely on techniques such as mutual information [27] to determine intensity correspondence relationships between two images. Hence, this complex phase representation allows for the use of more constrained similarity objective functions than mutual-information based objective functions. Further extensions to the complex wavelet phase representation from Eq. (3.11) to improve structural sensitivity and reduce noise sensitivity is discussed in Section 3.4 and Section 3.5 respectively.

3.4 Structural Sensitivity

A major challenge that needs to be addressed in the design of the complex phase representation is in dealing with the poor structural contrast inherent in images captured using

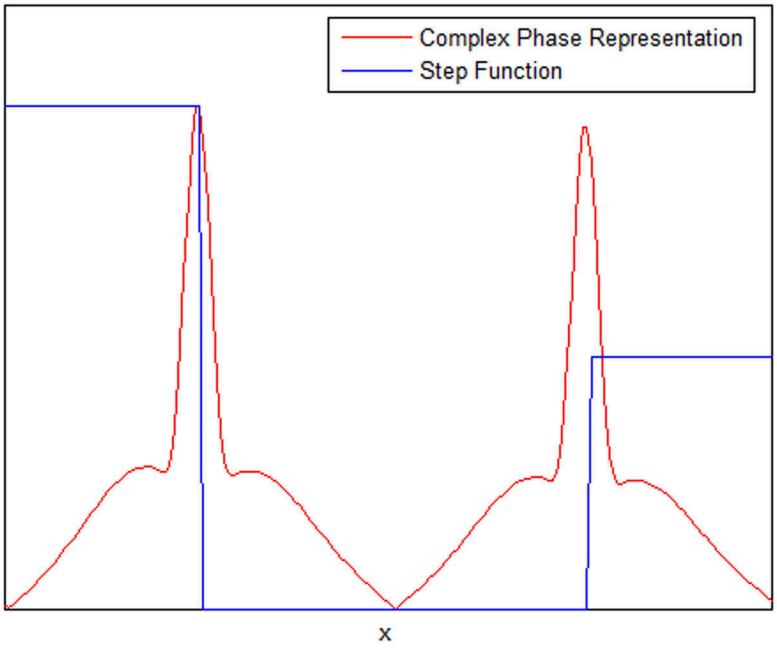

Figure 3.8: The complex wavelet phase representation $R(x)$ (Eq. (3.11)) of the piece-wise step function.

certain imaging systems such as fluoroscopy [68] and laser scanning ophthalmoscope [69]. Poor structural contrast makes it difficult to capture and distinguish fine structural detail, which is very important in quantifying the structural dissimilarities between two images. One approach to addressing the issues associated with poor structural contrast is to increase the response sensitivity of the complex wavelet representation to structural characteristics. Since the proposed complex phase representation relies on complex phase relationships to capture and distinguish structural characteristics, it is important to first study the behavior of the measure $R(\underline{x})$ described in Eq. (3.11) with respect to phase order. Let ϖ represent the phase deviation between a phase angle ϕ_θ and the amplitude-weighted mean phase angle $\bar{\phi}_\theta$,

$$\varpi = \phi_\theta - \bar{\phi}_\theta. \tag{3.12}$$

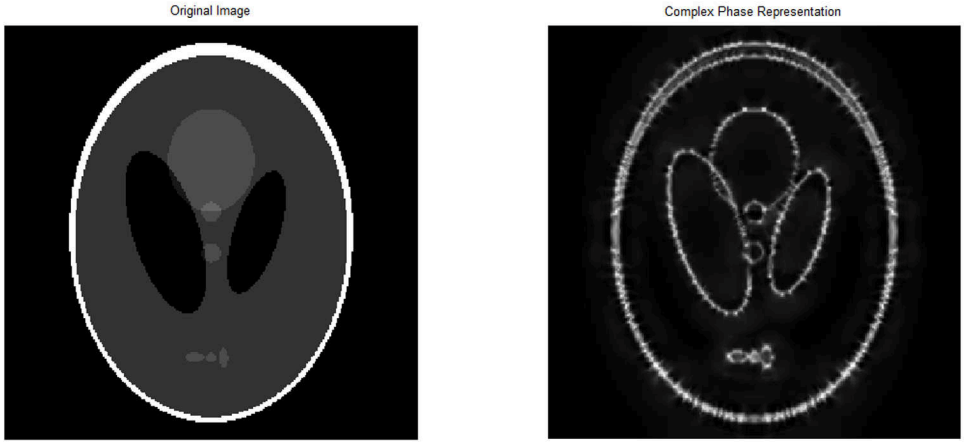

Figure 3.9: The complex wavelet phase representation $R(\underline{x})$ (Eq. (3.11)) of the Shepp-Logan phantom test image computed using a 2D Log-Gabor complex wavelet transform with four scales and six wavelet orientations. The structural information from the phantom test image are well preserved in the complex phase representation. Furthermore, the structural significance indicated by the complex phase representation is robust to contrast and intensity differences in the image, as can be seen by the boundaries between the white and black regions having similar structural significance as the boundaries between the light-gray and dark-gray regions.

Based on Eq. (3.11), R is proportional to a periodic phase deviation weighting function

$$\Lambda = \cos(\varpi). \tag{3.13}$$

As such, the analysis can be simplified by studying the behavior of Λ. It can be also observed in Eq. (3.11) that the absolute value of the phase deviation term $|\varpi|$ decreases as phase order increases, which corresponds to an increase in structural significance. Furthermore, given the properties of the cosine function, the decrease of $|\varpi|$ results in an increase of Λ. Therefore, one can study the response sensitivity of Λ to phase order by taking the derivative of Λ with respect to ϖ,

$$\frac{\partial \Lambda}{\partial \varpi} = -\sin(\varpi). \tag{3.14}$$

A plot of both $\Lambda = \cos(\varpi)$ and its derivative $\frac{\partial \Lambda}{\partial \varpi}$ is shown in Fig. 3.10a. Based on $\frac{\partial \Lambda}{\partial \varpi}$, the rate of change of Λ near the peak of Λ (where ϖ is minimal and phase order is maximal) is

very slow and gradual. For example, a reasonably large phase deviation of 37^o would still result in a large value of Λ that is 80% of the maximum value of Λ. This slow, gradual rate of change near the peak of Λ results in the poor response sensitivity of $R(\underline{x})$ to phase order and consequently to significant structural characteristics.

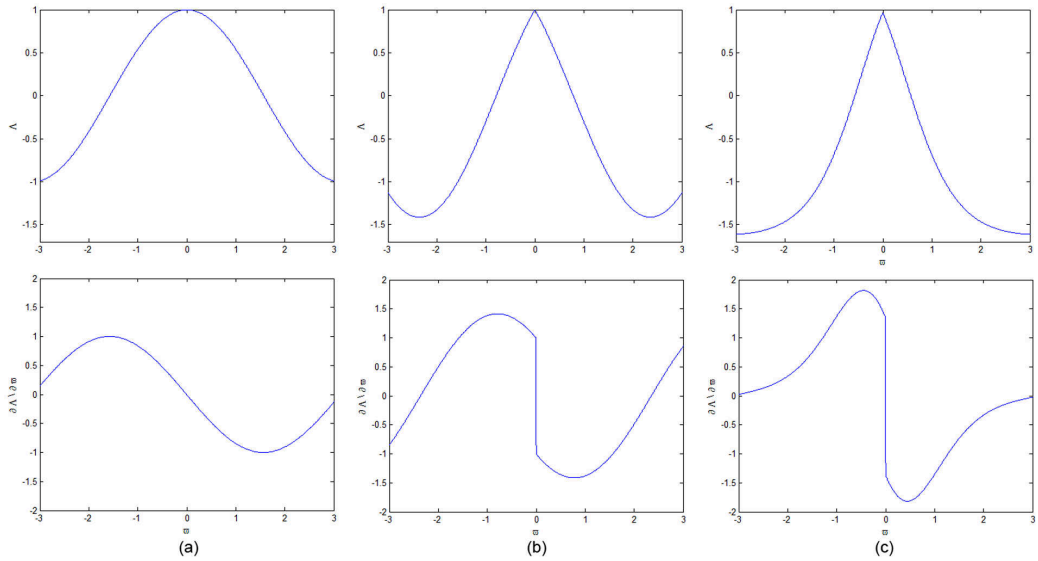

Figure 3.10: The plots of Λ (top) and its derivative $\frac{\partial \Lambda}{\partial \varpi}$ (bottom) for: a) weighting function proposed by Morrone et al. [55], b) weighting function proposed by Kovesi [65], and c) proposed weighting function (Eq. (3.15)). The response sensitivity of the proposed weighting function to significant structural characteristics is noticeably higher than the other weighting functions. It can also be observed that the proposed weighting function does not suffer from the undesirable phenomenon exhibited by the weighting function proposed by Kovesi [65], where the function Λ decreases as phase deviation $|\varpi|$ decreases in certain regions.

In an attempt to address this issue, Kovesi [65] proposed an alternative periodic phase deviation weighting function that exhibits a sharper rate of change near the peak of Λ, as shown in Fig. 3.10. While this weighting function increases the response sensitivity of $R(\underline{x})$ to significant structural characteristics when compared to the cosine weighting function proposed by Morrone et. al [55], there are two important limitations that need to be considered. First, the weighting function decreases almost linearly from minimum $|\varpi|$ to maximum $|\varpi|$. As such, the response sensitivity of the weighting function to phase order, and consequently to significant structural characteristics does not noticeably increase

as phase order increases. Hence, the weighting function does not emphasize structural characteristics with high structural significance, which is important in capturing and distinguishing fine structural detail. More importantly, the weighting function exhibits regions where the decrease of $|\varpi|$ actually results in a decrease of Λ, which is counterintuitive. To address all of the aforementioned limitations associated with the weighting function proposed by Kovesi [65], an alternative periodic weighting function based on the exponential difference of sinusoids is proposed,

$$\Lambda = \exp\left[\cos(\varpi) - \left|\sin\left(\frac{\varpi}{2}\right)\right|\right] - \frac{3}{2}. \tag{3.15}$$

Based on Eq. (3.15), the corresponding derivative $\frac{\partial \Lambda}{\partial \varpi}$ can be expressed as,

$$\frac{\partial \Lambda}{\partial \varpi} = \exp\left[\cos(\varpi) - \left|\sin\left(\frac{\varpi}{2}\right)\right|\right]\left(-\sin(\varpi) - \frac{1}{2}\frac{\partial\left|\sin\left(\frac{\varpi}{2}\right)\right|}{\partial \varpi}\cos\left(\frac{\varpi}{2}\right)\right). \tag{3.16}$$

The plots of the proposed weighting function (Eq. (3.15)) and its corresponding derivative (Eq. (3.16)) are shown in Fig. 3.10c. The proposed weighting function emphasizes high phase order (low $|\varpi|$), as indicated by the high rate of change near the peak of Λ. Furthermore, the proposed weighting function Λ never decreases as phase deviation $|\varpi|$ decreases, thus avoiding the undesirable phenomenon exhibited by the weighting function proposed by Kovesi [65]. Given this new weighting function Λ, the modified complex phase representation $R(\underline{x})$ can be expressed as,

$$R(\underline{x}) = \frac{\sum_{q=1}^{\beta}\sum_{s=1}^{\alpha} A_{s,\theta_q}(\underline{x})\left(\exp\left[\cos(\phi_{s,\theta_q}(\underline{x}) - \bar{\phi}_{\theta_q}(\underline{x})) - \left|\sin\left(\frac{\phi_{s,\theta_q}(\underline{x}) - \bar{\phi}_{\theta_q}(\underline{x})}{2}\right)\right|\right] - \frac{3}{2}\right)}{\sum_{q=1}^{\beta}\sum_{s=1}^{\alpha} A_{s,\theta_q}(\underline{x}) + e}, \tag{3.17}$$

where e is a small constant to handle situations where all the amplitude components are zero.

3.5 Noise Sensitivity

Another major challenge that needs to be addressed in the design of the complex phase representation is in dealing with image noise. Real-world images are also characterized by

noise, which makes it difficult to capture and distinguish structural detail with which to quantify the structural dissimilarities between images. As such, it is of great importance to study the effects of noise on the complex phase representation, as well as devise methods for reducing the noise sensitivity of the representation to produce meaningful image representations.

To study the effects of noise in the complex phase representation described in Eq. (3.17), additive Gaussian noise was applied with standard deviations of $\sigma = \{8\%, 16\%, 22\%, 25\%\}$ to a set of images acquired from the NLM Visible Human Project. A summary of the test images is given below.

1. **BMR**: Brain, axial, 1mm resolution, PD-weighted MR.
2. **PMR**: Pelvis, axial, 1mm resolution, T2-weighted MR.
3. **TMR**: Torso, coronal, 1.875mm resolution, T1-weighted MR.
4. **BCT**: Brain, axial, 1mm resolution, CT.
5. **PCT**: Pelvis, axial, 1mm resolution, CT.
6. **TCT**: Torso, coronal, 1.875mm resolution, CT.

The test images are shown in Fig. 3.11. To provide a quantitative assessment of the noise sensitivity of the complex phase representation, the peak signal-to-noise ratio (PSNR) and the mean Structural Similarity (MSSIM) [70] were measured between the complex phase representations with noise and without noise.

The PSNR and MSSIM results for the tested images are shown in Fig. 3.12 and Fig. 3.13 respectively. The PSNR is noticeably low under all noise scenarios for all test images. Furthermore, the MSSIM is very low for all noise scenarios as well. This means that the complex phase representation acquired under the noise scenarios provides a poor characterization of the structural information from the original image. The low PSNR and MSSIM results indicate that the complex phase representation described in Eq. (3.17) is highly sensitive to image noise, which is undesirable for providing robust feature representations with which to quantify structural dissimilarities between images. The complex phase representations of the BMR and BCT images under the tested noise scenarios are shown in Fig. 3.14 and Fig. 3.15. The complex phase representations are highly contaminated by noise under all noise scenarios. Furthermore, increasing structural degradation is exhibited in the complex phase representations as the level of noise increases. Therefore,

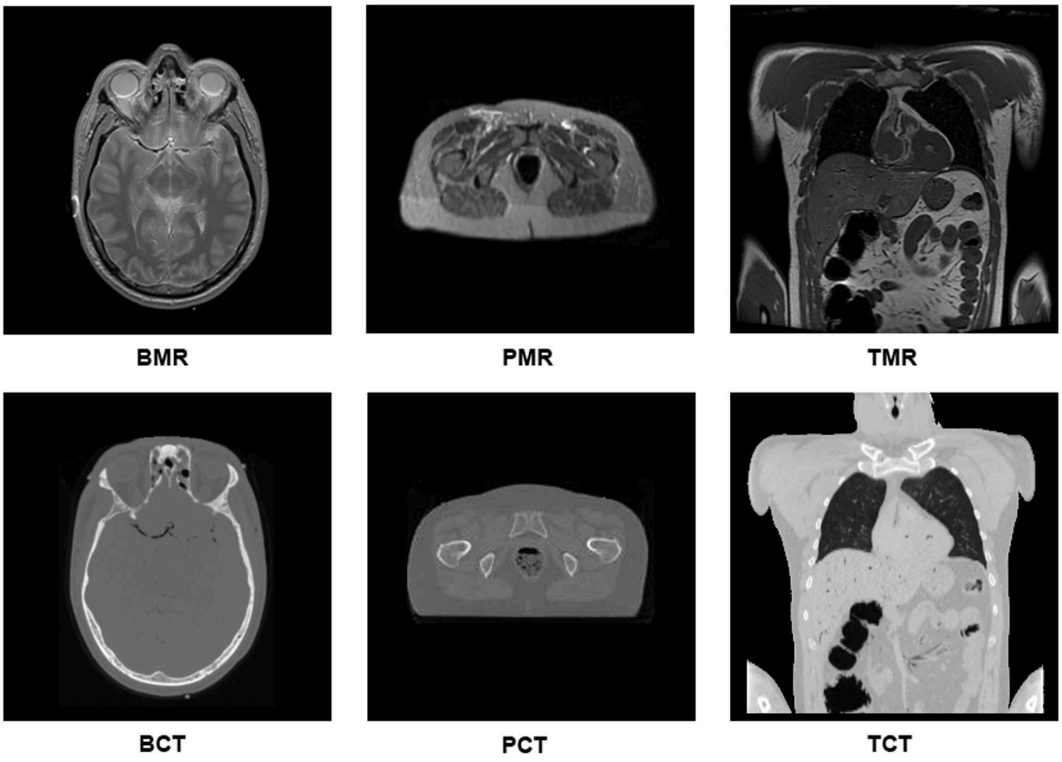

Figure 3.11: Test images used for noise sensitivity tests.

improving the robustness of the complex phase representation to image noise is necessary.

In an attempt to address this issue, Kovesi [65] proposed the use of a hard thresholding scheme to suppress noise characteristics in the constructed representation, where all values below the noise threshold are set to zero. While this hard thresholding approach decreases the sensitivity of $R(\underline{x})$ to noise, there are two main limitations to this approach. First, it is difficult to establish a hard noise threshold for noise suppression without losing important structural information in the constructed representation. Second, since all values below the hard noise threshold are set to zero, this results in a discontinuous feature representation, which is often undesirable. To address all of the aforementioned limitations associated with the weighting function proposed by Kovesi [65], the integration of a bilateral soft thresholding scheme that reduces the effect of noise in the complex phase representation while still providing a continuous representation for similarity comparison purposes is proposed.

The bilateral soft thresholding scheme can be described as follows. Let the summation

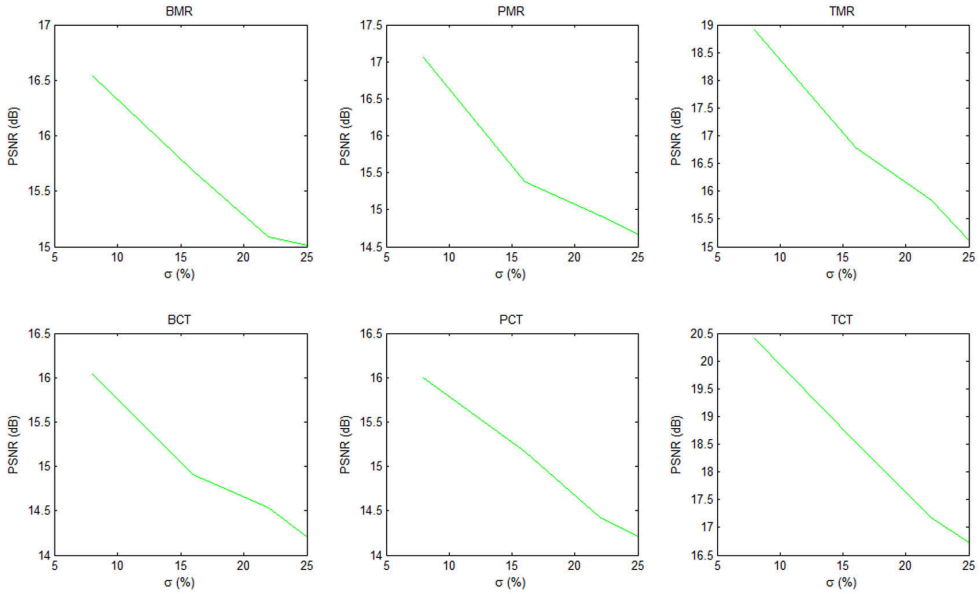

Figure 3.12: Plots of the PSNR for the tested images (Fig. 3.11) under different noise scenarios. The PSNR is noticeably low under all noise scenarios for all test images.

of complex phase deviations at point $\underline{x}$ be denoted by $\vartheta(\underline{x})$,

$$\vartheta(\underline{x}) = \sum_{q=1}^{\beta} \sum_{s=1}^{\alpha} A_{s,\theta_q}(\underline{x}) \Lambda(\underline{x}). \tag{3.18}$$

Two general observations can be made with regards to image noise with respect to complex phase and amplitude relationships. First, in uniform regions within the image where there is no structural characteristics, the complex phase deviations across all scales from the amplitude-weighted mean phase should be zero. Therefore, non-zero summations of complex phase deviations in these uniform regions are entirely due to the influence of noise and should be suppressed accordingly. Second, the distribution of complex amplitude across multiple scales is generally narrow and skewed towards high frequencies, whereas for strong structural characteristics the complex amplitudes are generally high and distributed evenly across scales [71]. Therefore, soft thresholding can be performed by weighting the summation of complex phase deviations based on the non-zero summations in uniform regions as well as the distribution of complex amplitude across scales. The bilateral weighting function W can therefore be expressed as a product of a weighting function $W_\eta(\underline{x})$ related

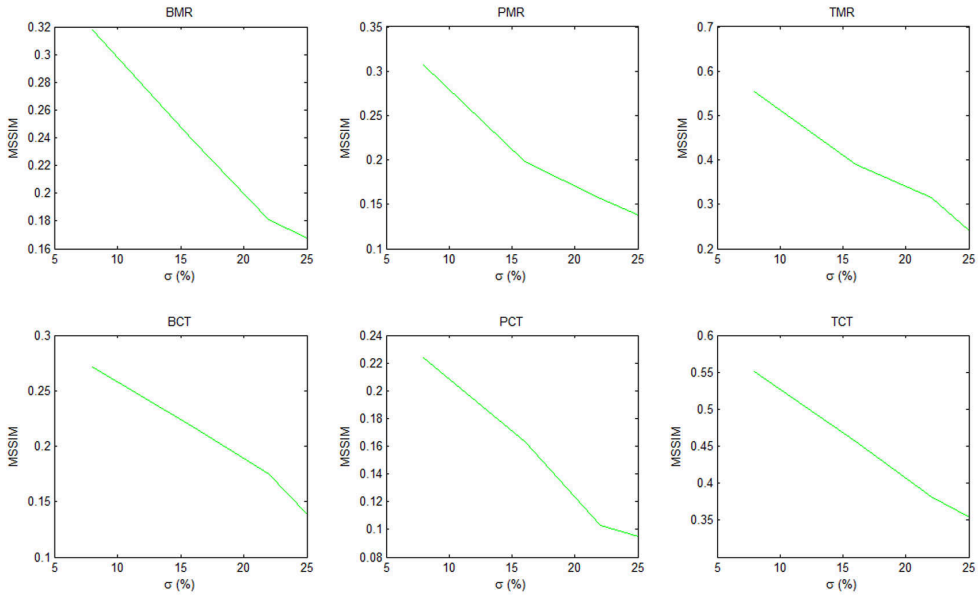

Figure 3.13: Plots of the MSSIM for the tested images (Fig. 3.11) under different noise scenarios. The MSSIM is very low for all noise scenarios as well. This means that the complex phase representation acquired under the noise scenarios provides a poor characterization of the structural information from the original image.

to non-zero summations in uniform regions and a weighting function $W_\tau(\underline{x})$ related to the distribution of amplitude across scales,

$$W(x,y) = W_\eta(x,y)W_\tau(x,y), \tag{3.19}$$

where $W_\eta(x,y)$ is governed by the median summation of phase deviations $\bar{\vartheta}$,

$$W_\eta(\underline{x}) = \frac{1}{1+\exp\left[\zeta(\bar{\vartheta}-\vartheta(\underline{x}))\right]}, \tag{3.20}$$

where ζ is the W_η decay factor, and $W_\tau(x,y)$ is governed by an estimate of how evenly distributed complex amplitude is across scales [65],

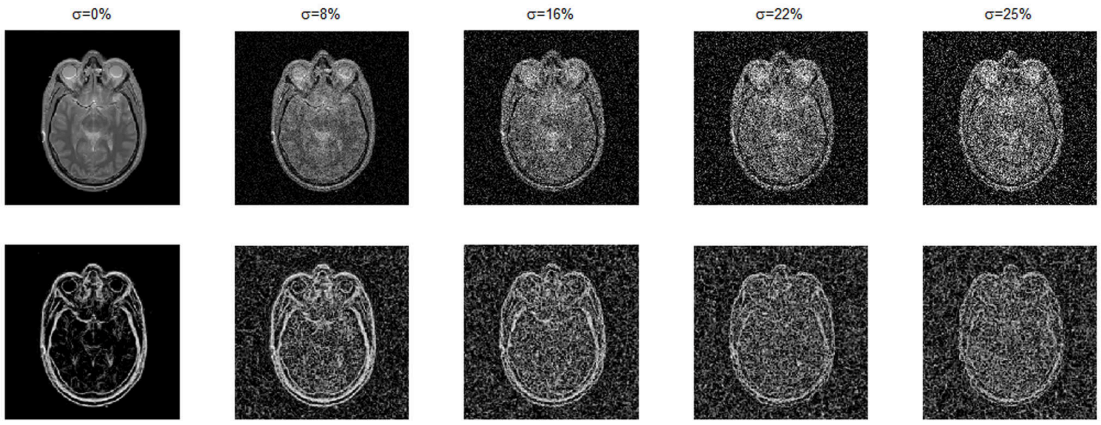

Figure 3.14: The complex phase representations (Eq. (3.17)) of the BMR test image (Fig. 3.11) under the tested noise scenarios. The top row shows the test image under the tested noise scenarios, while the bottom row shows the corresponding complex phase representations. The complex phase representations of BMR are highly contaminated by noise under all noise scenarios. Furthermore, increasing structural degradation is exhibited in the complex phase representations as the level of noise increases.

$$W_\tau(x,y) = \frac{1}{1+\exp\left(\epsilon_\tau c - \epsilon_\tau + \frac{\epsilon_\tau A_{max,\theta}(x,y)}{\sum_{s=1}^{\alpha} A_{s,\theta}(x,y)}\right)}. \tag{3.21}$$

where c is the frequency where $W_\tau = 0.5$ and ϵ_τ is the W_τ decay factor. Based on testing, setting $c = 0.4$ and $\epsilon = \zeta = 10$ was found to be effective and is used for all tests. Eq. (3.20) and Eq. (3.21) show that the penalty increases as the summation of complex phase deviations decreases to the range characterizing the uniform regions, as well as when the distribution of complex amplitude narrows respectively.

Using the bilateral soft thresholding scheme described in Eq. (3.19), the final complex phase representation can be expressed as,

$$R(\underline{x}) = \frac{\sum_{q=1}^{\beta}\sum_{s=1}^{\alpha} W_{\theta_q}(\underline{x}) A_{s,\theta_q}(\underline{x}) \Lambda(\underline{x})}{\sum_{q=1}^{\beta}\sum_{s=1}^{\alpha} A_{s,\theta_q}(\underline{x})}. \tag{3.22}$$

To study the effects of noise in the final complex phase representation with soft thresh-

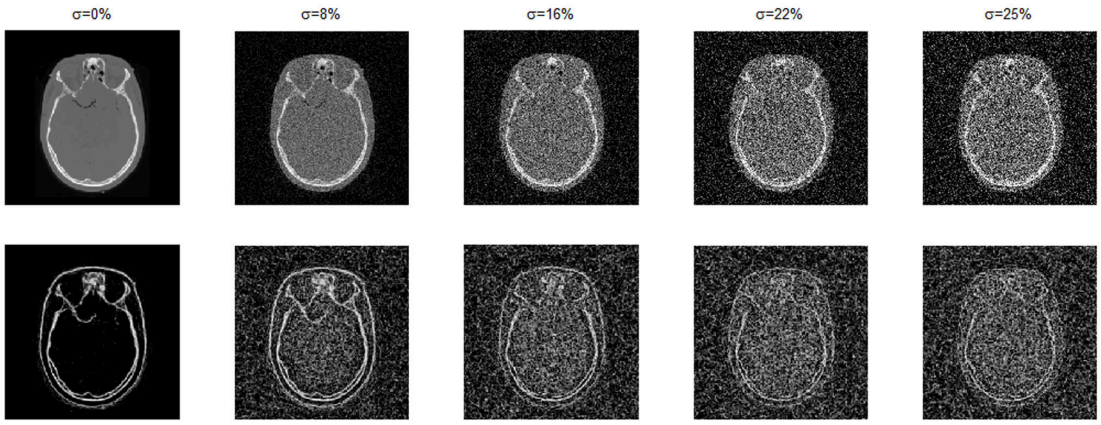

Figure 3.15: The complex phase representations of the BCT test image (Fig. 3.11) under the tested noise scenarios. The top row shows the test image under the tested noise scenarios, while the bottom row shows the corresponding complex phase representations. The complex phase representations of BCT are highly contaminated by noise under all noise scenarios. Furthermore, increasing structural degradation is exhibited in the complex phase representations as the level of noise increases.

olding described in Eq. (3.22), the same additive Gaussian noise test scenarios used to test the representation without soft thresholding described in Eq. (3.17) were applied to the new representation and the PSNR and MSSIM [70] were measured between the complex phase representations with noise and without noise.

The PSNR and MSSIM results for the tested images are shown in Fig. 4.8 and Fig. 4.9 respectively. The PSNR for the new representation with soft thresholding is significantly higher than without soft thresholding for all noise scenarios. Furthermore, the MSSIM for the representation with soft thresholding is significantly higher than without soft thresholding for all noise scenarios. What this means is that the complex phase representation with soft thresholding provides significantly improved characterization of the structural information from the original image than without soft thresholding. As such, the complex phase representation with soft thresholding described in Eq. (3.22) is robust to image noise, which is desirable for providing robust feature representations with which to quantify structural dissimilarities between images. The complex phase representations of the BMR and BCT images with soft thresholding under the tested noise scenarios are shown in Fig. 4.10. The complex phase representations are largely free of noise under all noise scenarios.

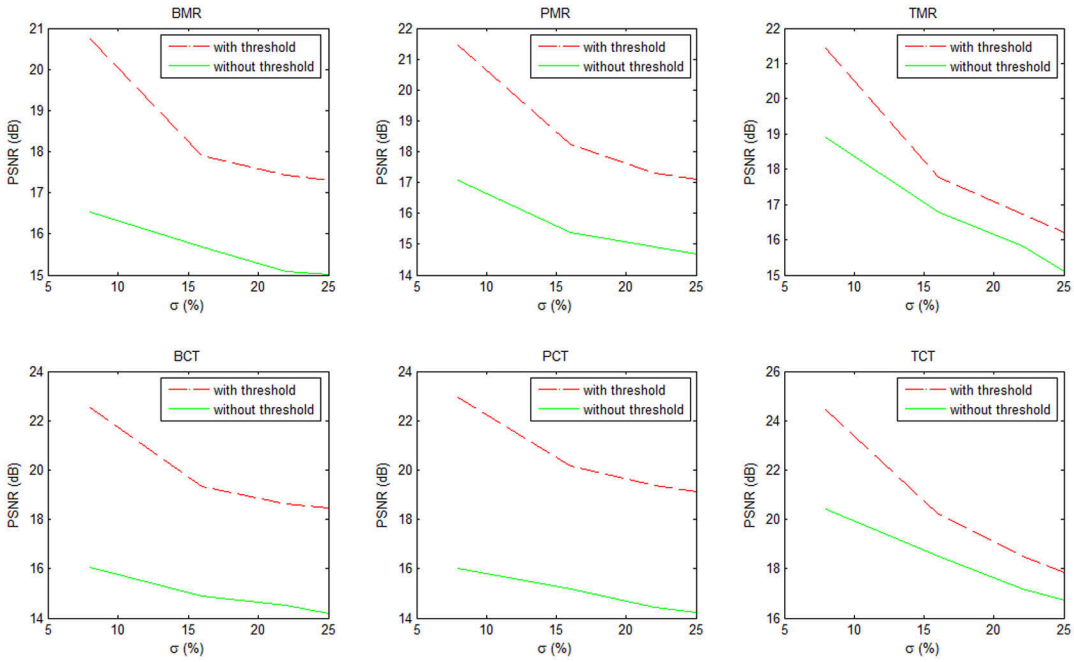

Figure 3.16: Plots of the PSNR for tested images (Fig. 3.11) under different noise scenarios with and without soft thresholding. The PSNR for the new representation with soft thresholding is significantly higher than without soft thresholding for all noise scenarios.

Unfortunately, the complex phase representation with soft thresholding suffers from two major limitations, as shown in Fig. 4.10. First, noticeable structural degradation is exhibited in the complex phase representations at high levels of noise, with representations acquired above $\sigma = 22\%$ largely unusable for the purpose of image registration due to significant structural degradation. This is due to the fact that under high levels of noise it is difficult to distinguish noise from image detail using the thresholding scheme, hence resulting in important structural characteristics being removed as well. Second, while the large-scale structural characteristics of the MR and CT images are similar, the MR image contain differing fine-scale structural characteristics than the CT image due to the properties of the imaging technologies. These two limitations motivate us to explore multi-scale approaches to producing feature representations for robust multimodal image registration even under high levels of noise and fine-scale structural differences between images, which we will discuss in the following chapter.

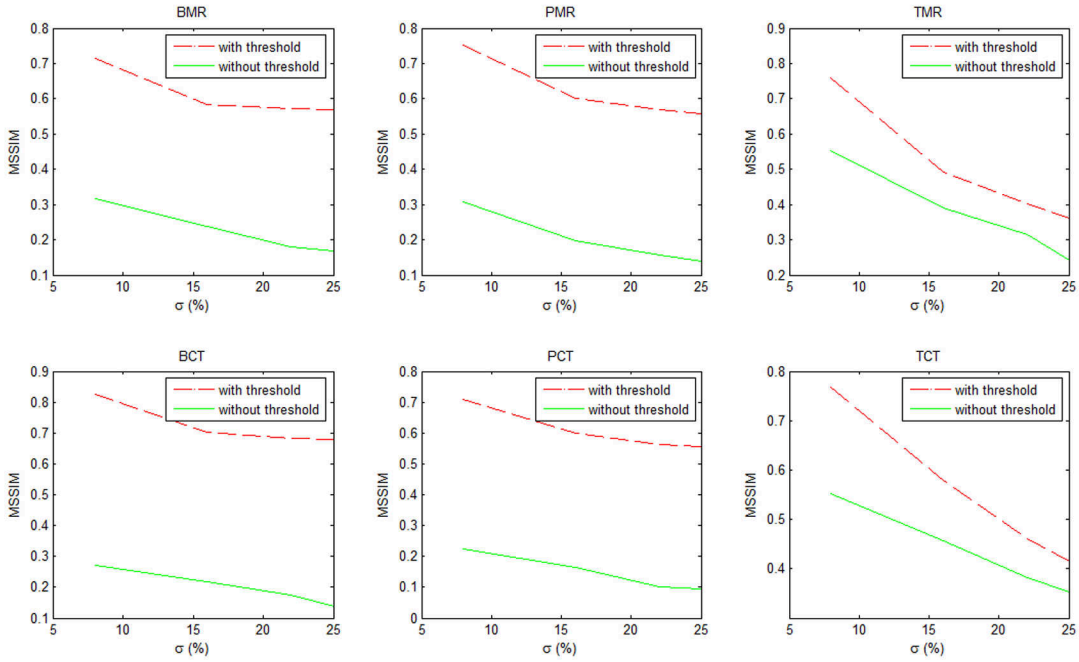

Figure 3.17: Plots of the MSSIM for tested images (Fig. 3.11) under different noise scenarios with and without soft thresholding. The MSSIM for the representation with soft thresholding is significantly higher than without soft thresholding for all noise scenarios. What this means is that the complex phase representation with soft thresholding provides significantly improved characterization of the structural information from the original image than without soft thresholding.

3.6 Summary

In this chapter, a novel complex wavelet representation based on phase relationships that extends upon [65] was presented. The proposed representation was designed to provide strong response to structurally significant characteristics, robust to contrast non-uniformities, and robust to noise. Unfortunately, based on a study of noise sensitivity, it was shown that noticeable structural degradation is exhibited in the complex phase representations at high levels of noise. Furthermore, the representation in its current form does not take into account fine-scale structural differences between images due to the underlying imaging modality. As such, we are motivated to extend the complex wavelet

representation in subsequent chapters to produce feature representations for robust multi-modal image registration even under high levels of noise and fine-scale structural differences between images.

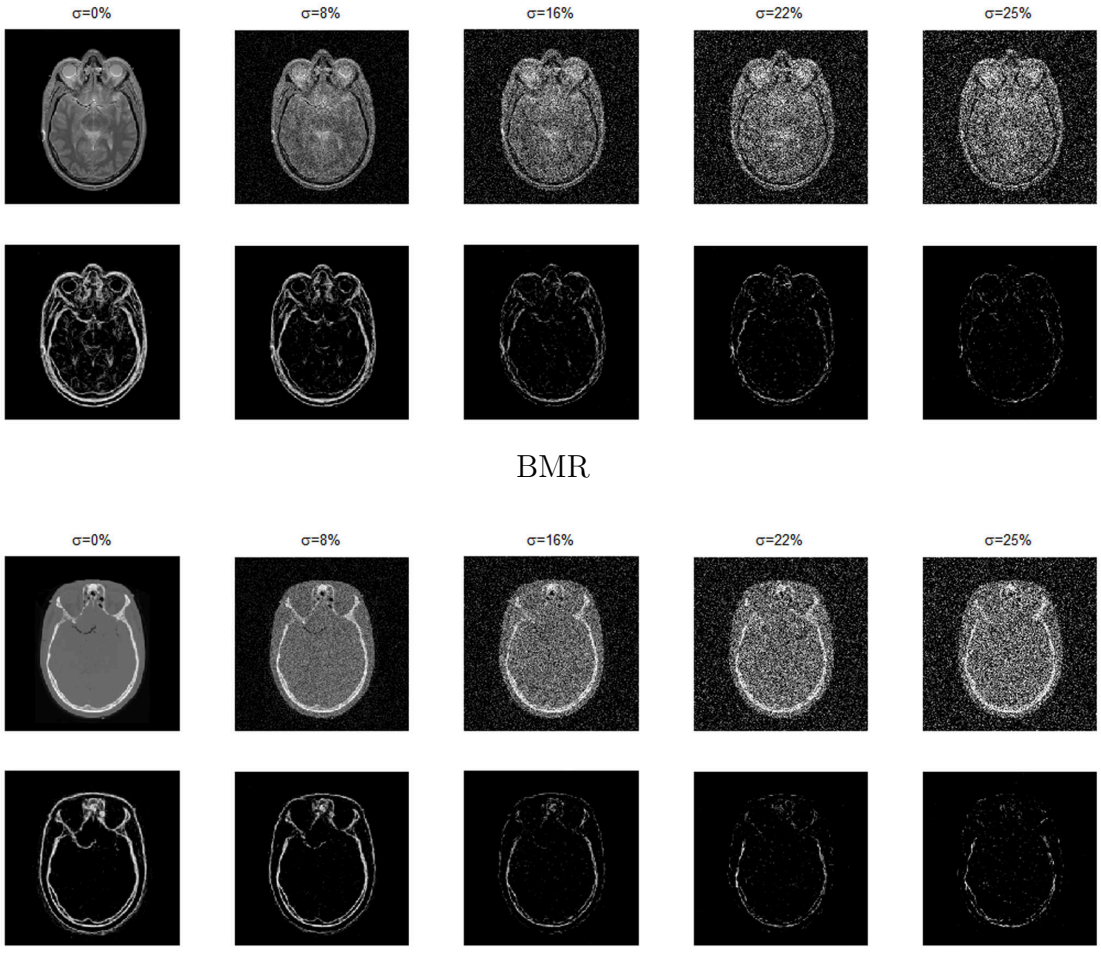

Figure 3.18: BMR and BCT images (Fig. 3.11) (top row of each panel) and the corresponding complex phase representations with soft thresholding (bottom row of each panel) under the tested noise scenarios. The complex phase representations of the BMR and BCT images are largely free of noise under all noise scenarios. Unfortunately, noticeable structural degradation is exhibited in the complex phase representations at high levels of noise, with representations acquired above $\sigma = 22\%$ largely unusable for the purpose of image registration due to significant structural degradation.

Chapter 4

Probabilistic Multi-Scale Representations

This chapter describes in detail the overall design and performance analysis of the proposed probabilistic complex phase representation. A introduction to scale space theory is presented. An overview of existing scale space representations is presented to examine the challenges and issues that need to be addressed in the design of multi-scale representations. A probabilistic approach to multi-scale complex phase representation is described in detail. Finally, the structural localization and noise sensitivity of the probabilistic complex phase representation are studied.

4.1 Introduction

There are two main limitations to the complex phase representation proposed in Chapter 3. First, while the complex phase representation is robust to noise, it suffers from significant structural degradation in situations characterized by low signal-to-noise ratio. As such, the resulting complex phase representation of images acquired under these high noise scenarios are poorly suited for the purpose of image registration, where the alignment between images is determined by comparing feature representations of the images. Second, images acquired using different imaging technologies can contain differing structural characteristics at various scales. For example, CT images provides an excellent characterization of large-scale bone structures, but does not capture fine-scale tissue structures. Conversely, MR images provide an excellent characterization of fine-scale tissue structures and a good

characterization of large-scale bone structures. Therefore, while the large-scale structural characteristics in MR and CT images are similar, the fine-scale structural characteristics can differ greatly, which can potentially affect the measure of similarity between the images. Therefore, registration inaccuracies may arise when using the complex phase representation proposed in Chapter 3 due to such fine-scale structural differences.

One approach to addressing the two main limitations associated with the complex phase representation is to extend it into a multi-scale representation through the use of scale space theory. Scale space theory is a framework for handling the inherent multi-scale nature of images by representing an image across multiple scales, with an increasing amount of fine-scale structures removed at each successive scale. One of the motivations for scale-space theory stems from the idea that, given no prior information about the scale of structures in an image, the only reasonable course of action is to represent the image at multiple pre-determined scales [72]. In scale space theory as first formalized by Witkin [73] and Koenderink and Van Doorn [74], an image $f(\underline{x})$ is represented as a single-parameter family of derived images $L(\underline{x};t)$, where t is a scaling parameter that defines the scale of details being represented. At each scale t, all structures in the image of smaller than a particular spatial size as governed by t are suppressed at the corresponding scale-space level. The scale-space representation $L(\underline{x};t)$ of $f(\underline{x})$ at a particular scale t can be alternatively formulated as the weighted integral over a local neighborhood ϱ around $\underline{x}$ in $f(\underline{x})$,

$$L(\underline{x};t) = \int_{\underline{a}\in\varrho} w(\underline{a},\underline{x};t) f(\underline{a})\, d\underline{a}, \tag{4.1}$$

where $w(\underline{a},\underline{x};t)$ is a weighting function that determines the contribution of $\underline{a}$ to the computation of $L(\underline{x})$ at scale t.

The use of scale-space representations allows image characteristics to be compared across multiple scales to reduce the effects of small-scale local geometric and noise distortions on the comparison process. For example, comparisons at fine scales are highly sensitive to local variations in position and orientation as well as noise, whereas comparisons at coarse scales are more robust to these local geometric and noise distortions. However, comparisons at fine scales allow for greater discrimination capabilities that help differentiate image characteristics that look similar at coarse scales. Therefore, a fine balance between image discrimination capabilities and robustness to local geometric and noise distortions can be achieved by taking multiple scales into account. Furthermore, the use of scale-space representations allows for greater robustness to fine-scale structural differences between images acquired using different imaging modalities during the comparison process.

4.2 Existing Scale Space Representations

Given the benefits of scale space theory for robust image representation, several scale space representations have been proposed over the years. Scale space representations can be generally divided into two main groups: i) linear, and ii) nonlinear.

4.2.1 Linear Scale Space Representations

In linear scale space representations, an image is decomposed into a family of derived images based on a linear scale space operator. The most common linear scale space representation is the Gaussian scale space representation first formalized by Witkin [73] and Koenderink and Van Doorn [74]. Given an image $f(\underline{x})$, the Gaussian scale space representation $L(\underline{x};t)$ is defined as the convolution of $L(\underline{x};0) = f(\underline{x})$ and a scale-dependent Gaussian weighting function w,

$$w(\underline{a};t) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi t}} \exp\left[-\frac{a_i^2}{2t}\right]. \tag{4.2}$$

Furthermore, several criteria for space space representations were also presented [74]:

Causality: A structure at a coarse scale must exist at a finer scale. New structures must not be introduced going from a finer scale to a coarser scale.

Homogeneity and Isotropy: The scale space operator is required to be space and scale invariant.

The causality criterion guarantees that fine-scale structures are removed in a monotonic fashion as the scale increases, which is very important for clear discrimination between structures of different scales. This causality criterion is generally considered the most important criterion of scale space theory. On the other hand, the homogeneity and isotropy criterion was simply introduced by Koenderink and Van Doorn for the sake of analytical simplicity [74] and is generally not considered a requirement of scale space theory. The main advantage of Gaussian scale space representations is its computational and theoretical simplicity. Furthermore, the properties and characteristics of linear Gaussian scale space theory has been well-studied [75, 76], with alternative discrete Gaussian scale space representations being proposed that provided improved computational advantages over simple sampled Gaussian scale space representations [77].

The Gaussian scale space representations (Eq. (4.2)) of the PD-weighted MR and CT images used in Section 3.5 are shown in Fig. 4.1. There are two observations that can be made that illustrate the fundamental limitation of linear scale space representations. First, the large-scale structures at coarse scales are poorly localized when compared to their true locations in the original image, where well-localized structures refer to sharp structures that coincide with the true locations of the structures in the image. Second, much of the corner information exhibited in the original images is removed at coarse scales. This makes extracting meaningful structural information at coarse scales for comparison purposes during the image registration process very challenging. The significant structural degradation exhibited in the Gaussian scale space representations is largely due to the fact that the Gaussian scale space operator diffuses pixel intensities in an isotropic fashion, regardless of the underlying image content. As such, pixel intensities from different structures are combined together, thus violating the boundary between the structures and resulting in the poor structure localization and corner destruction. An additional limitation of the Gaussian scale space representation, which will be illustrated in Section 4.6, is its sensitivity to high levels of noise at fine scales, which is one of the issues with the complex phase representation. This sensitivity to high levels of noise is due to the fact that the local information redundancy upon which the Gaussian scale space operators rely on is insufficient to provide a good scale space estimates of the original image content under such situations. The noise sensitivity of Gaussian scale space representations will further studied in Section 4.6.

4.2.2 Nonlinear Scale Space Representations

To address the limitations of linear scale space representations, a new class of nonlinear scale space representations was introduced [78–82]. In nonlinear scale space representations, an image is decomposed into a family of derived images based on a nonlinear scale space operator. Given that the limitations of the linear scale space representations stem largely from the structure delocalizing and distorting nature of isotropic diffusion, a majority of existing nonlinear scale space representations extend from the isotropic diffusion in a non-linear fashion to better preserve meaningful structural information at coarse scales. The generalized diffusion equation can be expressed as

$$\frac{\partial L}{\partial t} = c(\underline{x}; t)\nabla^2 L + \nabla \cdot (c \nabla L), \tag{4.3}$$

where $c(\underline{x}; t)$ is the diffusion coefficient and ∇ is the gradient. In the case of isotropic

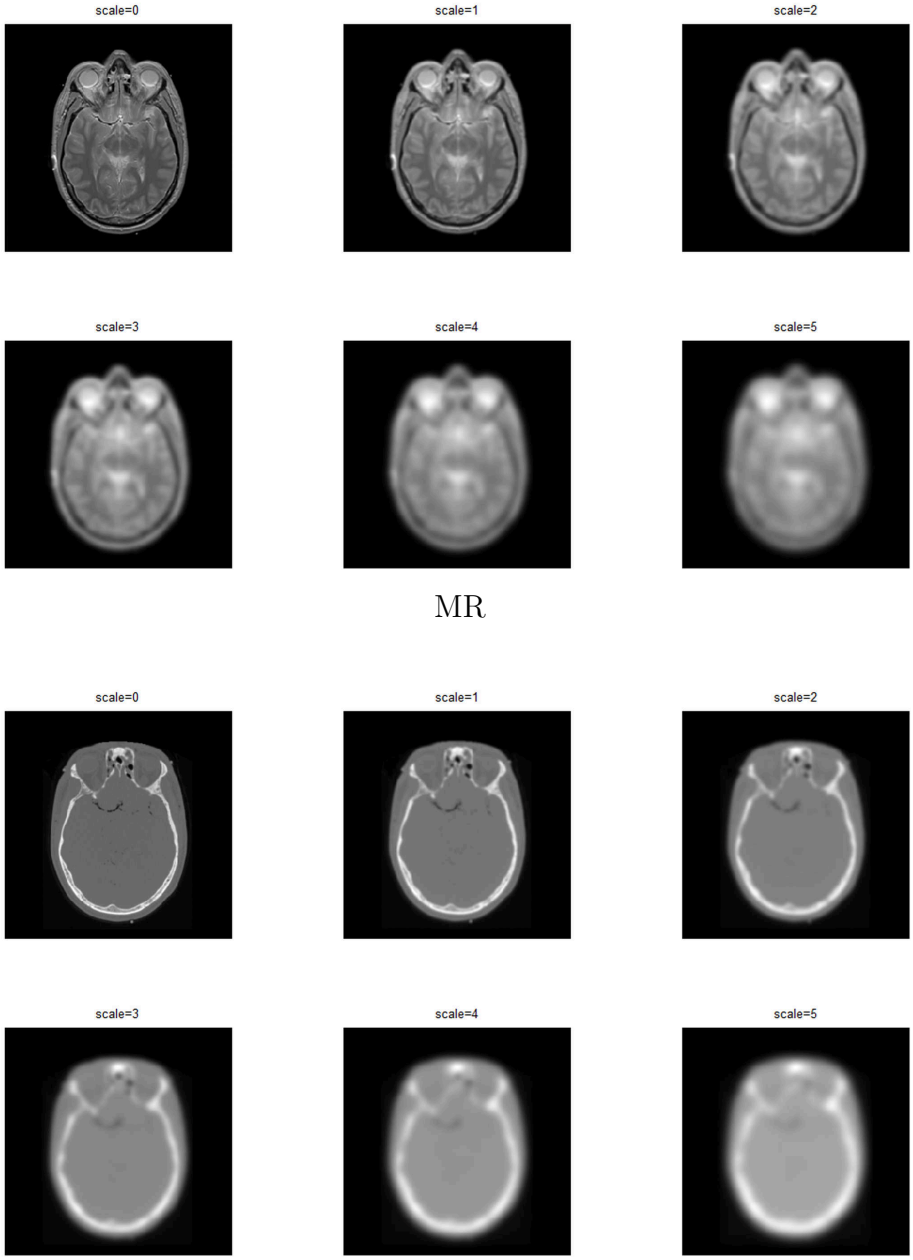

Figure 4.1: Gaussian scale space representation (Eq. (4.2)) of PD-weighted MR and CT axial brain images at different scales. The large-scale structures at coarse scales are poorly localized when compared to their true locations in the original images.

diffusion, $c(x;t)$ is set to a constant.

One of the most popular nonlinear scale-space representations is that proposed by Perona and Malik [78], where they make the claim that there was no reason why the homogeneity and isotropy criterion introduced by Koenderink and Van Doorn [74] was a necessary requirement of scale space theory. Furthermore, Perona and Malik reason that, in addition to the causality criterion, two other important criteria must be considered to obtain meaningful multi-scale space space representations of images [78]:

Immediate Localization: The structures at each scale should be sharp and coincide with the true locations of the structures in the image.

Piecewise Smoothing: Intraregion smoothing should have preference over interregion smoothing at all scales.

Both of these criterions indicate that structural characteristics should be well preserved at their corresponding scales, which is not satisfied by the structure delocalizing and distorting nature of isotropic diffusion. In an attempt to satisfy these criteria, Perona and Malik proposed the use of a non-negative diffusion coefficient in Eq. (4.3) that is a function of the gradient magnitude,

$$c(\underline{x};t) = h\left(\|\nabla L(\underline{x};t)\|\right). \tag{4.4}$$

In the 2D case, Perona and Malik proposed the use of the following diffusion coefficient,

$$c(x,y;t) = \exp\left[-\left(\frac{\nabla L(x,y;t)}{\kappa}\right)^2\right], \tag{4.5}$$

where κ is the decay constant. This diffusion coefficient $c(x,y;t)$ possesses two important behavioral characteristics that are important for satisfying the scale space criteria described by Perona and Malik. First, the diffusion coefficient $c(x,y;t)$ preserves structures by limiting diffusion when the gradient $\nabla L(x,y;t)$ is large, thus satisfying the immediate localization criterion. Second, the diffusion coefficient $c(x,y;t)$ promotes intraregion smoothing by allowing for greater diffusion in directions of lower gradient, thus satisfying the piecewise smoothing criterion. The discrete nonlinear diffusion-based scale space representation as described by Perona and Malik for 2D cases can be expressed as

$$L(x,y;t) = L(x,y;t-1) + \lambda \left(\sum_i \sum_j c(x+i, y+j; t) \nabla L(x+i, y+j; t-1) \right), \quad (4.6)$$

$$L(x,y;0) = f(x,y), \quad (4.7)$$

The main advantage of nonlinear diffusion-based scale space representations such as that presented in Eq. (4.6) is that they provide better structural preservation at different scales, which is important for providing meaningful structural information for comparison purposes during the image registration process.

The nonlinear diffusion-based scale space representations of the PD-weighted MR and CT images used in Section 3.5 are shown in Fig. 4.2. The structures are noticeably better preserved in the nonlinear diffusion-based scale space representations at coarse scales when compared to the Gaussian scale space representations. Unfortunately, the structures at the coarser scales start to become delocalized and the corner information is removed as with the case of Gaussian scale space representations, thus violating the immediate localization criterion at coarser scales. This structural degradation at coarser scales is largely due to the fact that nonlinear diffusion relies entirely on local information redundancy, which is insufficient for maintaining structural detail at the coarser scales. Furthermore, like the Gaussian scale space representation, nonlinear diffusion-based scale space representations are sensitive to high levels of noise at fine scales, which is one of the issues with the complex phase representation and is also attributed to the fact that the local information redundancy upon which the nonlinear diffusion-based scale space operators rely on is insufficient to provide a good scale space estimates of the original image content under such situations. The noise sensitivity of the nonlinear diffusion-based scale space representations will be studied in Section 4.6.

Given the limitations of current linear and nonlinear scale space representations under high noise scenarios, one is motivated to add the following criterion for scale space representations:

Noise Robustness: The resulting scale space representation should not be influenced by the presence of noise at all scales.

Furthermore, given that the two major limitations of nonlinear diffusion-based scale space representations are both associated with the limitations of local information redundancy, one is motivated to explore alternative nonlinear multi-scale representations that are able to utilize global information redundancy from the entire image to provide better structural preservation and localization at coarse scales and greater robustness to noise.

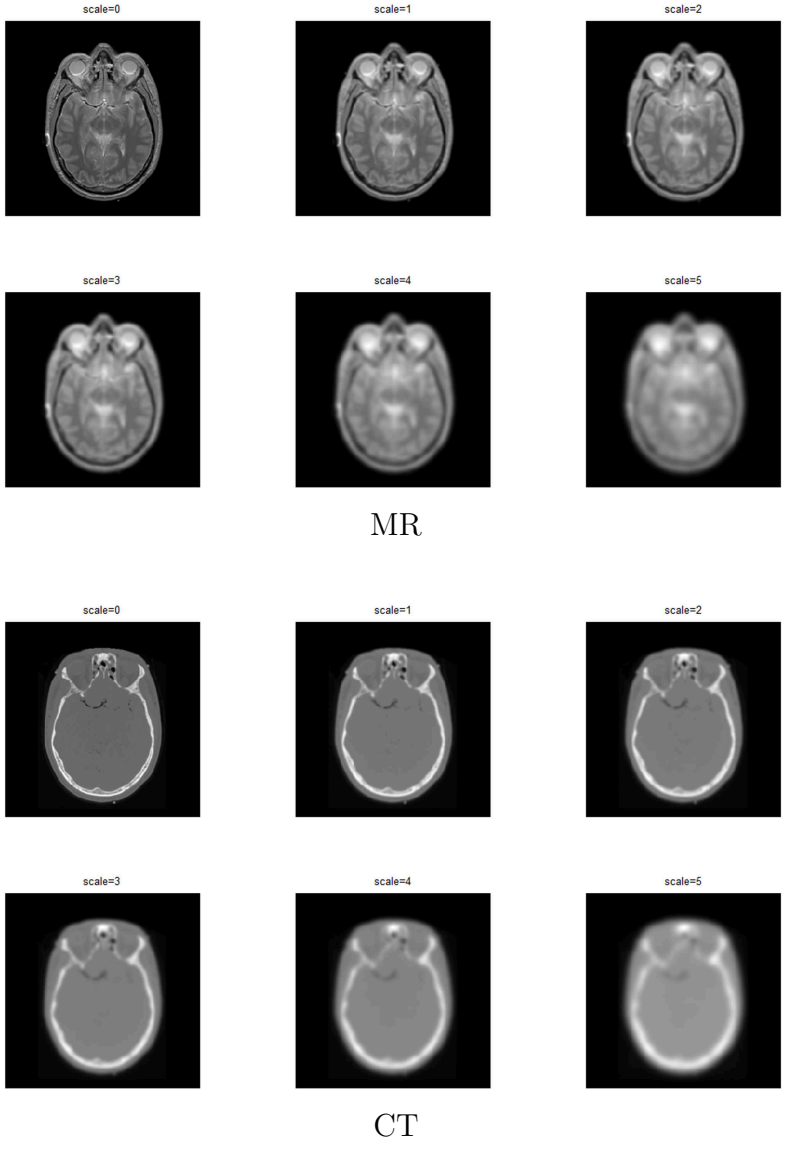

Figure 4.2: Nonlinear diffusion-based scale space representation (Eq. (4.6) of PD-weighted MR and CT axial brain images at different scales. The structures are noticeably better preserved in the nonlinear diffusion-based scale space representations at coarse scales when compared to the Gaussian scale space representations (Fig. 4.1). Unfortunately, the structures at the coarser scales starts to become delocalized and the corner information is removed as with the case of Gaussian scale space representations, thus violating the immediate localization criterion at coarser scales.

4.3 Probabilistic Multi-Scale Representation

The underlying assumption behind scale space theory is that structural characteristics of an image, as well as noise, are monotonically reduced by averaging an increasing number of samples together to provide a scale space estimate of the image $f(\underline{x})$. Based on this assumption, scale space representations rely on information redundancy within the image itself to reduce the visibility of image structures and noise. When taken in the context of the generalized scale space formulation in Eq. (4.1), nonlinear scale space representations (e.g., Eq. (4.6)) employ nonlinear weighting functions for $w(\underline{a}, \underline{x}; t)$ that adapt the contribution of a based on $f(\underline{x})$. There are two main limitations of existing nonlinear scale space representations. First, structures become delocalized and degraded at the coarser scales since the local information redundancy is insufficient for maintaining structural detail at the coarser scales. Second, existing nonlinear scale space representations generally perform poorly in situations characterized by low signal-to-noise ratios since the local information redundancy is insufficient to provide good scale space estimates of the original image content. Given that both issues associated with existing nonlinear scale space representations are due to the limitations of local information redundancy, one is motivated to extend the concept of nonlinear scale space representation to utilize global information redundancy by considering the space F of all pixels within $f(\underline{x})$ when computing the nonlinear scale space representation $L(\underline{x}; t)$,

$$L(\underline{x}; t) = \int_{\underline{a} \in F} w(\underline{a}, \underline{x}; t) f(\underline{a}) \, d\underline{a}. \tag{4.8}$$

Unfortunately, computing $L(\underline{x}; t)$ based on Eq. (4.8) is computationally infeasible in a deterministic manner. To address this computational issue, an alternative approach is proposed to instead compute a probabilistic estimate of $L(\underline{x}; t)$ as described in Eq. (4.8) using a probabilistic sampling approach. Consider a random field S representing all possible pixels within $f(\underline{x})$. If we were to take m random samples $\underline{\xi}_1, \ldots, \underline{\xi}_m$ from a distribution p, the probabilistic estimate of $L(\underline{x}; t)$ can be computed as

$$\hat{L}(\underline{x}; t) = \sum_{i=1}^{m} w(\underline{\xi}_i, \underline{x}; t) f(\underline{\xi}_i). \tag{4.9}$$

The main advantage of the proposed probabilistic approach to multi-scale representation shown in Eq. (4.9) is that it is much more computationally efficient to compute than the

extended nonlinear scale space representation described in Eq. (4.8), given that a relatively small sample set is sufficient for attaining similar benefits of global information redundancy.

There are two main issues that need to be addressed in constructing the probabilistic multi-scale representation as described in Eq. (4.9). First, the choice of distribution p used for drawing samples from is very important for obtaining a consistent estimate of representation $L(\underline{x}; t)$. A poorly chosen p can result in high estimation variance and inconsistent estimates. Second, the design of a nonlinear weighting function $w(\underline{\xi}_i, \underline{x}; t)$ that preserves large-scale structural localization while suppressing noise and fine-scale structures at coarser scales is needed. To address the issue of sampling distribution, an adaptive rejection sampling strategy is proposed to obtain a relevant sample set for constructing the probabilistic representation in a consistent manner, as described in Section 4.3.1. To address the nonlinear weighting function design issue, an adaptive weighting scheme is proposed based on the neighborhood gradient between $\underline{\xi}_i$ and $\underline{x}$.

4.3.1 Adaptive Rejection Sampling

As discussed in Section 4.3, one issue faced in the construction of probabilistic representations is the choice of a distribution p from which to draw random samples. A poorly chosen distribution can result in too many irrelevant samples being generated from S that contribute little useful information to the construction of $L(\underline{x}; t)$. Not only does this increase unnecessary computational overhead, but can also lead to undesirably high estimation variance. Therefore, a choice of p that yields the most relevant samples to the estimation of $L(\underline{x}; t)$ is desired to reduce estimation variance and improve computational efficiency.

Unfortunately, the distribution p that provides a consistent estimate of scale space representation $L(\underline{x}; t)$ is unknown and can vary greatly depending on the underlying image content. However, such a distribution should satisfy certain conditions. First, since the underlying goal of probabilistic multi-scale representation is to reduce structures and noise in a successive manner using global information redundancy, intuitively, the samples drawn from S based on p should have high information redundancy with $\underline{x}$. Second, structures with high information redundancy are typically within relatively close spatial proximity to each other. Therefore, the samples drawn from S based on p should be more likely within relatively close spatial proximity to $\underline{x}$. Based on these two conditions that the distribution p should satisfy, an adaptive rejection sampling strategy that allows samples to be drawn from such a distribution is proposed.

The proposed adaptive rejection sampling strategy can be described as follows. To draw

a random sample $\underline{\xi}$ from a distribution p, a sample $\underline{\chi}$ is first drawn from an exponential instrumental distribution q that adapts based on the spatial proximity and statistical similarity between $\underline{x}$ and all other sites within $f(\underline{x})$,

$$q(\underline{\chi}|\underline{x}) = \frac{\exp\left[-\left|\underline{x}-\underline{\chi}\right|^2\left(\sigma^2(\underline{\chi})-\sigma^2(\underline{x})\right)^2\right]}{\sum\limits_{\underline{\chi}}\exp\left[-\left|\underline{x}-\underline{\chi}\right|^2\left(\sigma^2(\underline{\chi})-\sigma^2(\underline{x})\right)^2\right]}, \tag{4.10}$$

where $\sigma^2(\cdot)$ is the local spatial variance and $|\cdot|$ is the Euclidean spatial distance between $\underline{x}$ and $\underline{\chi}$. The adaptive instrumental distribution described in Eq. (4.10) promotes the drawing of samples that are within close spatial proximity and have high statistical similarity, which are more likely to share high information redundancy with $\underline{x}$.

While the instrumental distribution q promotes samples with higher statistical similarity and close spatial proximity to $\underline{x}$, there is no guarantee that the samples have high information redundancy. To address this issue, the information redundancy between the sample $\underline{\chi}$ drawn from the instrumental distribution q and $\underline{x}$ is then evaluated based on the cumulative Gaussian-weighted squared gradient Φ between the local neighborhoods $\varrho_{\underline{x}}$ and $\varrho_{\underline{\chi}}$,

$$\Phi(\varrho_{\underline{x}}, \varrho_{\underline{\chi}}) = \sum_{\varrho} w_g(\varrho)\cdot\left(f(\varrho_{\underline{\chi}}) - f(\varrho_{\underline{x}})\right)^2. \tag{4.11}$$

where w_g is the Gaussian weighting function. Based on Eq. (4.11), the sample $\underline{\chi}$ is accepted or rejected based on the following condition,

$$\Phi(\varrho_{\underline{x}}, \varrho_{\underline{\chi}}) < \tau. \tag{4.12}$$

where τ is the rejection threshold. The rejection threshold τ should be adapted based on the underlying image content to better preserve structures as well as suppress noise. For images with low noise, more irrelevant samples (smaller τ) should be pruned, however for images characterized by high levels of noise, there is a greater tolerance for sample variation, and therefore an increased τ to preserve more samples. To accommodate for these variations, τ is adjusted based on an estimate of noise variance,

$$\tau = \frac{1}{2}{\sigma_n}^2. \tag{4.13}$$

If the condition in Eq. (4.12) holds, then the sample $\underline{\chi}$ is accepted as a realization of p. If the condition does not hold, then $\underline{\chi}$ is rejected. The rejection sampling process is repeated until an upper bound for sampling q has been reached, resulting in the final set of m samples $\underline{\xi}_1, \ldots, \underline{\xi}_m$ for reconstructing $L(\underline{x}; t)$.

4.3.2 Adaptive Weighting Function

As discussed in Section 4.3, the second issue faced in the construction of probabilistic multi-scale representations is the design of a nonlinear weighting function $w(\underline{\xi}_i, \underline{x}; t)$. The nonlinear weighting function should achieve three important goals. First, the weighting function should maintain structural localization and avoid structural degradation at coarse scales so that the immediate localization and piecewise smoothing criteria are met. Second, the weighting function should provide good effective noise suppression such that the resulting scale space representation is not influenced by the presence of noise at all scales, thus satisfying the noise robustness criterion. Third, the weighting function should remove structural information from the image in a monotonic fashion as the scale increases. By introducing no new structures going from a finer scale to a coarse scale, the causality criterion is satisfied. As such, the nonlinear weighting function used in the construction of the probabilistic representation needs to be designed to achieve all three goals.

Recall that to maintain structural localization and avoid structural degradation at coarse scales while suppressing noise at fine scales, a nonnegative exponential diffusion coefficient was introduced in the nonlinear diffusion-based scale space representations proposed by Perona and Malik [78],

$$c(x, y; t) = \exp\left[-\left(\frac{\nabla L(x, y; t)}{\kappa}\right)^2\right], \tag{4.14}$$

where κ is the decay constant. This diffusion coefficient limits diffusion when the gradient is large and promotes diffusion in directions of lower gradient, thus preserving structural characteristics within an image. Interpreting the diffusion coefficient as a nonlinear weighting function, the contribution of a sample $\underline{\xi}$ to the construction of $L(\underline{x}; t)$ is proportional to the image gradient. Inspired by the diffusion coefficient proposed by Perona and Malik, a nonlinear weighting function $w(\underline{\xi}, \underline{x})$ based on the concept of squared neighborhood gradient ∇_ϱ is introduced,

$$w(\underline{\xi}, \underline{x}) = \exp\left[-\left(\frac{\nabla_{\varrho}(\underline{\xi}, \underline{x})}{\kappa}\right)\right], \tag{4.15}$$

where κ is the decay constant and the neighborhood squared gradient ∇_{ϱ} is defined as the cumulative Gaussian-weighted squared gradient between the local neighborhoods $\underline{\varrho}_{\underline{x}}$ and $\underline{\varrho}_{\underline{\chi}}$ as defined in Eq. (4.11). Finally, to accommodate for monotonic structural reduction as the scale increases, we replace the decay constant κ with a diffusion function $\kappa(t)$ based on the scaling parameter t,

$$\kappa(t) = \eta t. \tag{4.16}$$

where η is the diffusion multiplier constant. Based on $\kappa(t)$, the final nonlinear weighting function $w(\underline{\xi}_i, \underline{x}; t)$ can be expressed as

$$w(\underline{\xi}, \underline{x}; t) = \exp\left[-\left(\frac{\nabla_{\varrho}(\underline{\xi}, \underline{x})}{\eta t}\right)\right]. \tag{4.17}$$

4.4 Probabilistic Complex Phase Representation

Based on the m samples $\underline{\xi}_1, \ldots, \underline{\xi}_m$ drawn using the adaptive rejection sampling scheme introduced in Section 4.3.1 and the nonlinear weighting function $w(\underline{\xi}_i, \underline{x}; t)$ introduced in Section 4.3.2, the probabilistic representation $L(\underline{x}; t)$ of an image $f(\underline{x})$ can be computed as

$$\hat{L}(\underline{x}; t) = \sum_{i=1}^{m} w(\underline{\xi}_i, \underline{x}; t) f(\underline{\xi}_i). \tag{4.18}$$

The probabilistic representations of the PD-weighted MR and CT images used in Section 3.5 are shown in Fig. 4.3, with the upper bound for sampling p set to a maximum of 100 iterations. The structures are significantly better preserved and localized in the probabilistic multi-scale representations at coarse scales when compared to the Gaussian and nonlinear diffusion-based scale space representations, thus better satisfying the immediate localization criterion at coarser scales. This is most noticeable in the CT image, where the large scale bone structures are well preserved and localized at all scales. This improved structural preservation and localization at coarser scales is largely due to the fact that the probabilistic multi-scale representation relies on the global information within an image,

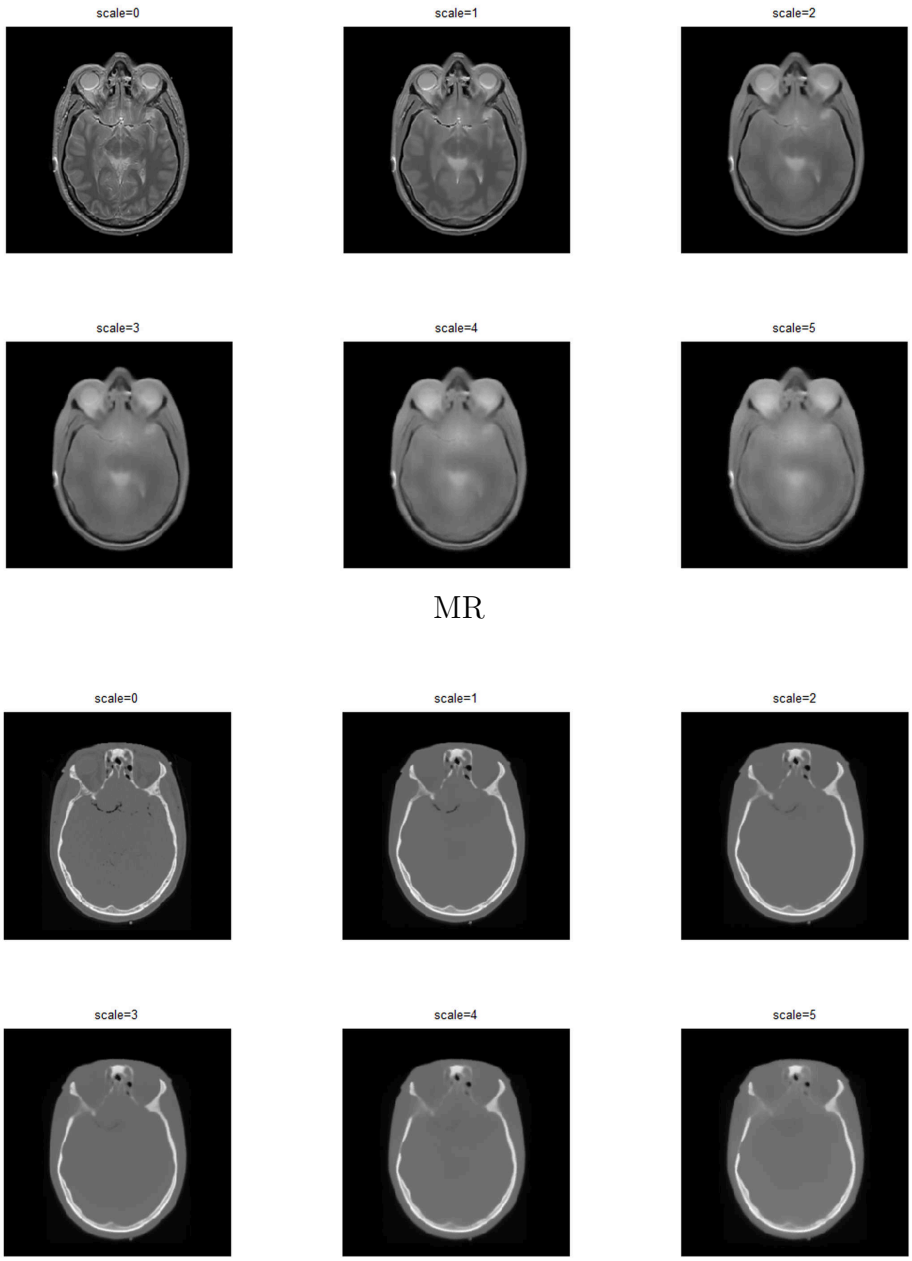

Figure 4.3: Probabilistic representation (Eq. (4.18)) of PD-weighted MR and CT axial brain images at different scales. The structures are significantly better preserved and localized in the probabilistic representations at coarse scales when compared to the Gaussian (Fig. 4.1) and nonlinear diffusion-based (Fig. 4.2) scale space representations.

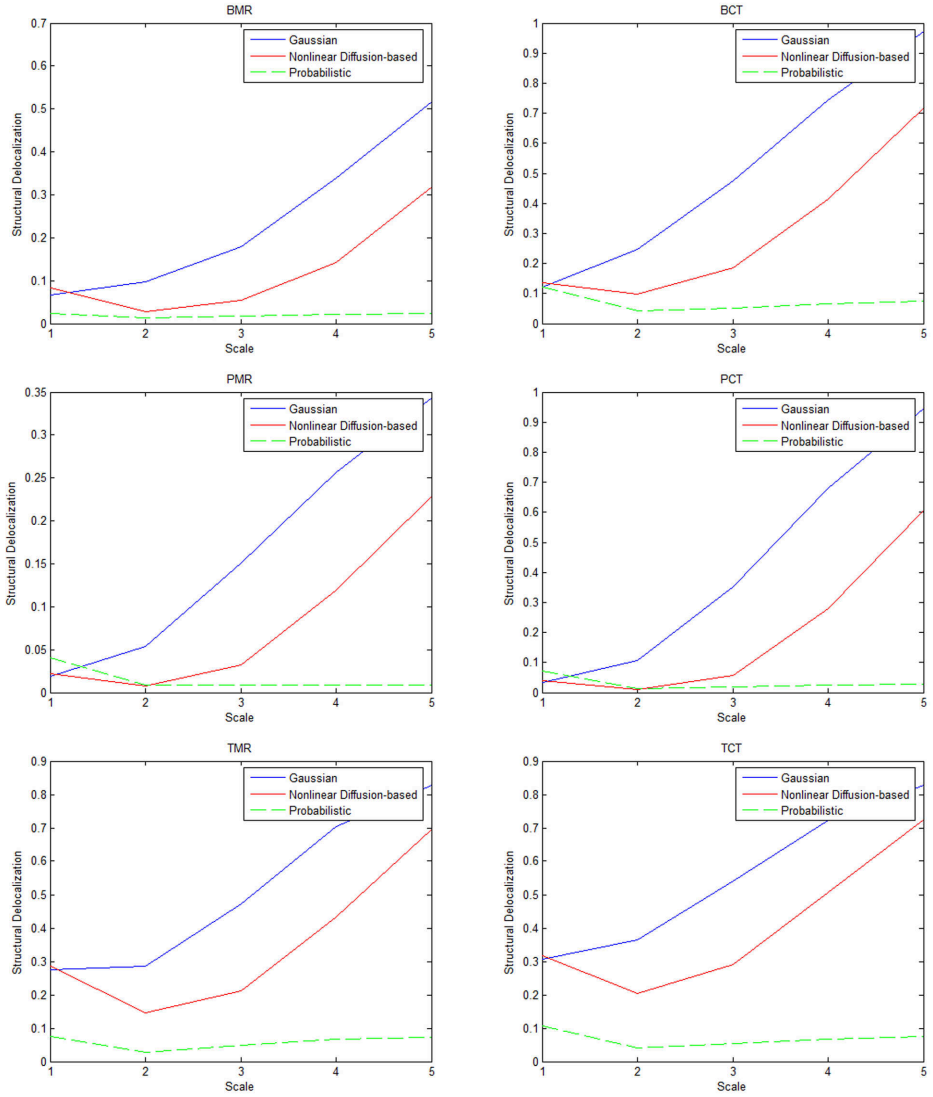

Figure 4.4: Plots of the structural delocalization ϵ_ν of the complex phase representations for the tested images (Fig. 3.11) at different scales. The structural delocalization of the complex phase representations constructed using linear Gaussian scale space and nonlinear diffusion-based scale space increases significantly as scale increase. On the other hand, the structural delocalization of the probabilistic complex phase representations remains relatively low and constant at all scales for all test cases.

which provides greater information redundancy to better maintain structural detail at the coarser scales.

Given that the underlying goal is to extend the complex phase representation proposed in Chapter 3 into a multi-scale representation through the use of the proposed probabilistic theory, the final probabilistic complex phase representation can be constructed by computing the complex phase representation $R(\underline{x})$ (Eq. (3.22)) of $L(\underline{x};t)$ at each scale t,

$$R(\underline{x};t) = \frac{\sum_{q=1}^{\beta}\sum_{s=1}^{\alpha} W_{\theta_q}(\underline{x};t)A_{s,\theta_q}(\underline{x};t)\left(\exp\left[\cos(\phi_{s,\theta_q}(\underline{x};t)-\bar{\phi}_{\theta_q}(\underline{x};t))-\left|\sin\left(\frac{\phi_{s,\theta_q}(\underline{x};t)-\bar{\phi}_{\theta_q}(\underline{x};t)}{2}\right)\right|\right]-\frac{3}{2}\right)}{\sum_{q=1}^{\beta}\sum_{s=1}^{\alpha} A_{s,\theta_q}(\underline{x};t)}, \quad (4.19)$$

where $A_{s,\theta}(\underline{x};t)$ and $\phi_{s,\theta}(\underline{x};t)$ is computed for $L(\underline{x};t)$ at each scale t according to Eq. (3.4) and Eq. (3.5) respectively,

$$A_{s,\theta}(\underline{x};t) = \sqrt{\left(L(\underline{x};t) * J^e_{s,\theta}(\underline{x})\right)^2 + \left(L(\underline{x};t) * J^o_{s,\theta}(\underline{x})\right)^2}, \quad (4.20)$$

and,

$$\phi_{s,\theta}(\underline{x};t) = \tan^{-1}\left(\frac{L(\underline{x};t) * J^e_{s,\theta}(\underline{x})}{L(\underline{x};t) * J^o_{s,\theta}(\underline{x})}\right). \quad (4.21)$$

An algorithmic summary of the proposed probabilistic complex phase representation is presented in Algorithm 1.

Algorithm 1 Probabilistic complex phase representation

1: **for** $t = 1$ to NUMSCALES **do**
2: Compute probabilistic multi-scale representation L_f and L_g (4.9) for f and g.
3: Compute complex amplitude A_f, A_g (4.20) and phase ϕ_f, ϕ_g (4.21) for L_f and L_g.
4: Compute complex phase representation R_f, R_g from (A_f,ϕ_f) and (A_g,ϕ_g) (4.19).
5: **end for**

By constructing the multi-scale complex phase representation based on the probabilistic framework, structural characteristics should be better localized and preserved at coarse scales, while attaining robustness to scenarios characterized by low signal-to-noise ratios at fine scales. To validate claims of immediate localization, the structural localization of the probabilistic complex phase representation will be studied in Section 4.5. Furthermore, to validate claims of noise robustness, the noise sensitivity of the probabilistic complex phase representation will be studied in Section 4.6.

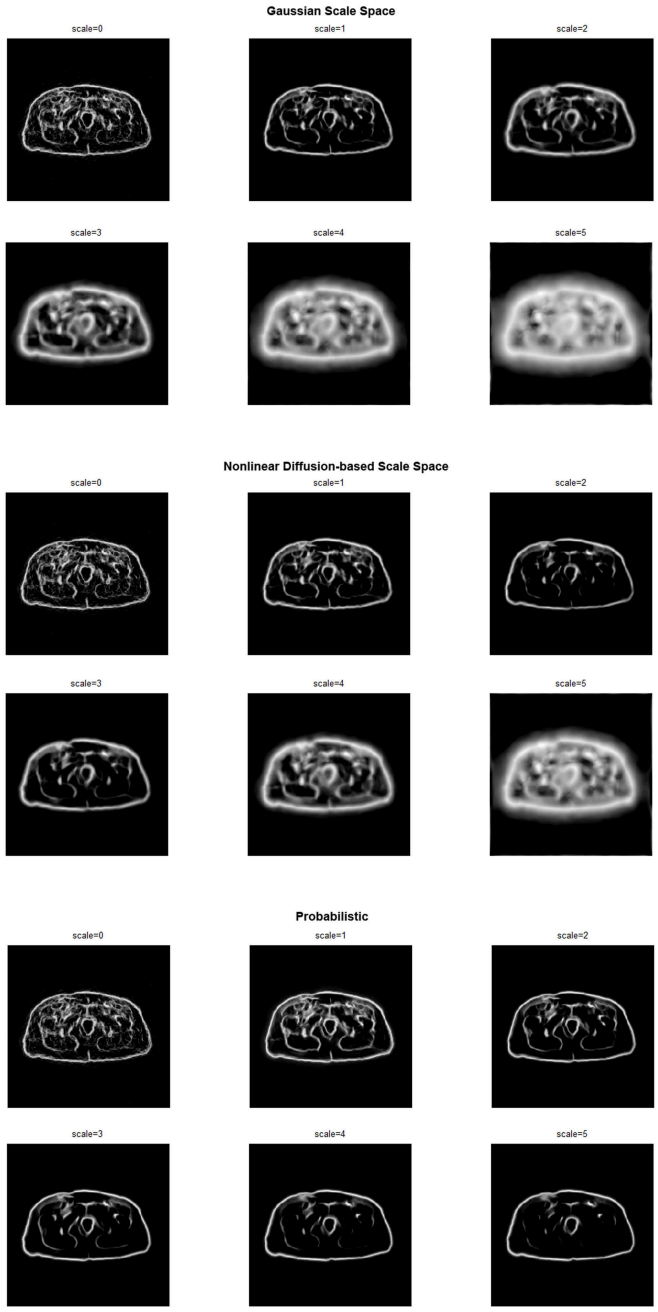

Figure 4.5: The multi-scale complex phase representations of the PMR test image at different scales constructed using Gaussian and nonlinear diffusion-based scale spaces, and the proposed probabilistic framework. The structural characteristics in the image are significantly better localized in the probabilistic complex phase representations at coarse scales when compared to the Gaussian and nonlinear diffusion-based scale space representations.

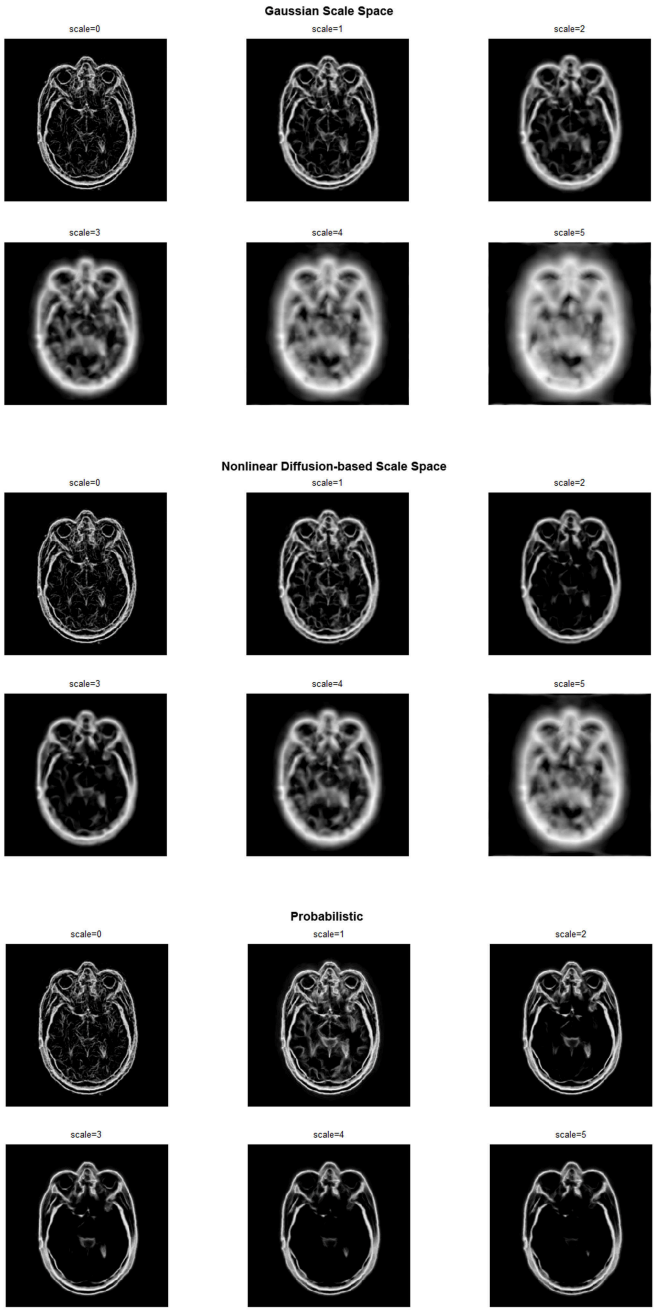

Figure 4.6: The multi-scale complex phase representations of the BMR test image at different scales constructed using Gaussian and nonlinear diffusion-based scale spaces, and the proposed probabilistic framework. The fine scale tissue structures are reduced monotonically while large scale structures remain localized as scale increases.

4.5 Structural Localization

To validate that the probabilistic complex phase representation described in Eq. (4.19) satisfies the immediate localization criterion, studying the localization of structures at different scales is important. This is accomplished by constructing the probabilistic complex phase representations for a set of real medical image images acquired from the NLM Visible Human Project described in Section 3.5. For comparison purposes, scale space complex phase representations using linear Gaussian scale space and nonlinear diffusion-based scale space were also constructed.

Given the constructed complex phase representations, designing a measure of structural localization is necessary in order to perform a quantitative analysis between the different multi-scale representations. According to the immediate localization criterion, the locations of structures at any given scale should also coincide with their locations in the original image. Therefore, intuitively, all structural information at a particular scale that does not exist within the original image is considered a result of structural delocalization. Motivated by this, an effective measure of structural delocalization, ϵ, for the multi-scale representation at a particular scale t can be defined as the cumulative structural significance of structures that do not coincide with existing structures in the original image,

$$\epsilon\left(t\right) = \sum_{x, R_{\upsilon}(x;0)=0} R_{\upsilon}(x;t). \tag{4.22}$$

One issue with this measure of structural delocalization is that it does not take into account the quantity of actual structural characteristics within the original image. To address this issue, one should divide the structural delocalization measure in Eq. (4.22) by the quantity of structures in the original image,

$$\epsilon_{\nu}\left(t\right) = \frac{\epsilon\left(t\right)}{\sum\limits_{x}\{R_{\upsilon}(x;0) > 0\}}. \tag{4.23}$$

The structural delocalization ϵ_{ν} of the complex phase representations for the tested images at different scales t are shown in Fig. 4.4. The structural delocalization of the complex phase representations constructed using linear Gaussian scale space and nonlinear diffusion-based scale space increases significantly as scale increase. On the other hand, the structural delocalization of the complex phase representations constructed using the probabilistic framework remains relatively low and constant at all scales for all test cases.

What this means is that the probabilistic complex phase representation provides good structural localization at all scales, thus better satisfying the immediate localization criterion than the linear Gaussian and nonlinear diffusion-based scale space complex phase representations.

The multi-scale complex phase representations of the PMR, BMR, and TCT test images at different scales are shown in Fig. 4.5, Fig. 4.6, and Fig. 4.7 respectively. The structural characteristics in the image are significantly better localized in the probabilistic complex phase representations at coarse scales when compared to the Gaussian and nonlinear diffusion-based scale space representations, thus better satisfying the immediate localization criterion at coarser scales. This is most noticeable in the BMR test case, where the fine scale tissue structures are reduced monotonically while large scale structures remain localized as scale increases. This improved structural localization at coarser scales is largely due to the fact that the probabilistic multi-scale representation relies on the global information within an image, which provides greater information redundancy to better preserve the true location of structures at the coarser scales.

4.6 Noise Sensitivity

To study the effects of noise on the probabilistic complex phase representation described in Eq. (4.19), additive Gaussian noise was applied with standard deviations of $\sigma = \{8\%, 16\%, 22\%, 25\%\}$ of the dynamic range to the test images used in Section 3.5. To provide a quantitative assessment of the noise sensitivity of the probabilistic complex phase representation at fine scales, the peak signal-to-noise ratio (PSNR) and the mean Structural Similarity (MSSIM) [70] were measured between the complex phase representations at scale $t = 1$ with noise and without noise. For comparison purposes, the noise sensitivity tests were also performed on multi-scale complex phase representations constructed using Gaussian scale space and nonlinear diffusion-based scale space. The noise sensitivity analysis was performed at scale $t = 1$ for two main reasons. First, the influence of noise on multi-scale representations are most prominent at fine scales, thus making it representative of the noise sensitivity of a multi-scale representation. Second, the structural localization at that scale remains similar across all multi-scale representations, thus allowing for a fair comparison of noise sensitivity between the different multi-space representations. The PSNR and MSSIM results for the complex phase representation described in Section 3.5 is shown as a baseline reference.

The PSNR and MSSIM results for the tested images are shown in Fig. 4.8 and Fig. 4.9

respectively. The PSNR and MSSIM are noticeably higher for the multi-scale representations for all noise scenarios than the baseline reference complex phase representation, thus illustrating the gain in noise robustness when scale space representations are used. Furthermore, the PSNR and MSSIM of the probabilistic complex phase representations is noticeably higher than the Gaussian and nonlinear diffusion-based representations. In particular, the probabilistic complex phase representation exhibited PSNR gains of approximately 2 dB at $\sigma = 25\%$ in the BMR, BCT, and PCT cases and MSSIM gains greater than 0.1 at $\sigma = 25\%$ in the BMR and BCT cases. What this means is that the probabilistic complex phase representation provides an improved characterization of the structural information from the original image under high levels of noise than the other tested representations. The multi-scale complex phase representations of the BMR image under the tested noise scenarios using Gaussian scale space, nonlinear diffusion-based scale space, and probabilistic framework are shown in Fig. 4.10. While increasing structural degradation is exhibited in all of the tested multi-scale complex phase representations as the level of noise increases, the probabilistic complex phase representations are able to better preserve structural detail at high noise levels when compared to the other tested multi-scale complex phase representations. Therefore, the proposed probabilistic complex phase representation is able to better satisfy the noise robustness criterion.

4.7 Summary

In this chapter, a novel probabilistic complex phase representation was presented. The proposed multi-scale representation was designed to address the issues associated with robustness to high levels of noise and fine-scale structural differences between images exhibited by the complex phase representation described in Chapter 3. A study of structural localization demonstrated that the proposed representation provides improved structural localization at coarse scales when compared to linear and nonlinear diffusion-based scale space complex phase representations. Furthermore, a study of noise sensitivity showed that the proposed representation provides improved structural detail preservations at high noise levels when compared to the other tested multi-scale complex phase representations. Given the proposed robust feature representation, an image registration method that utilizes this representation is needed for determining the alignment between images in an automated manner.

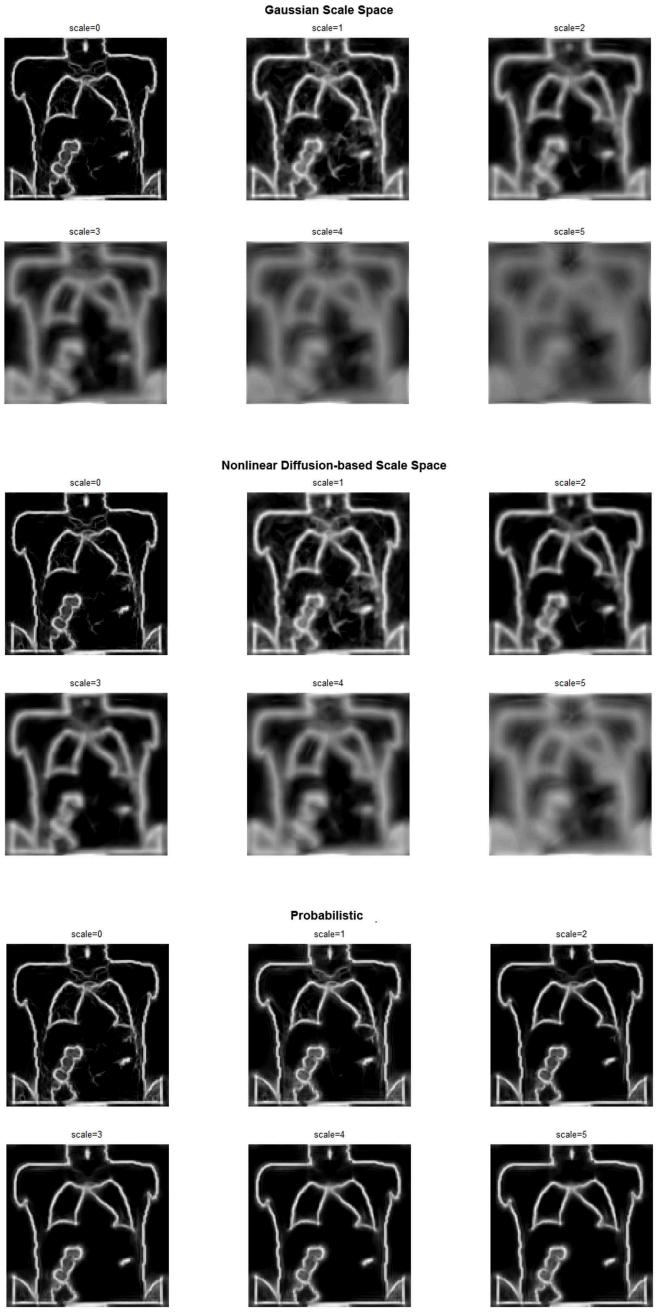

Figure 4.7: The multi-scale complex phase representations of the TCT test image at different scales constructed using Gaussian and nonlinear diffusion-based scale spaces, and the proposed probabilistic framework.

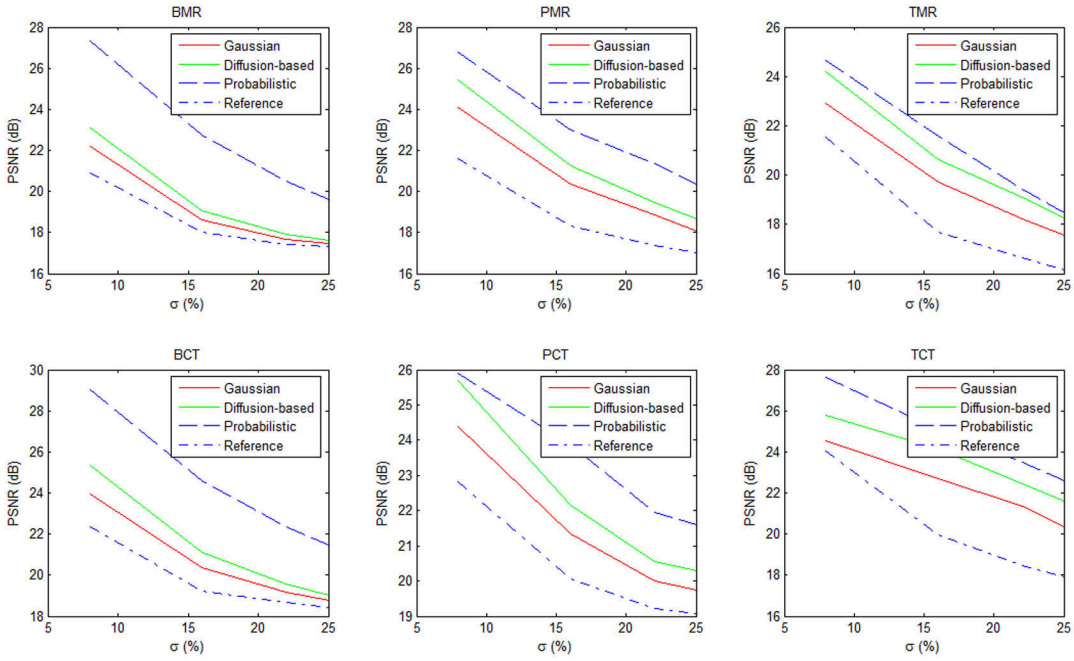

Figure 4.8: Plots of the PSNR for the tested images (Fig. 3.11) under different noise scenarios. The PSNR is noticeably higher for the multi-scale representations for all noise scenarios than the baseline reference complex phase representation, thus illustrating the gain in noise robustness when multi-scale representations are used. Furthermore, the PSNR of the probabilistic complex phase representations is noticeably higher than the Gaussian and nonlinear diffusion-based representations.

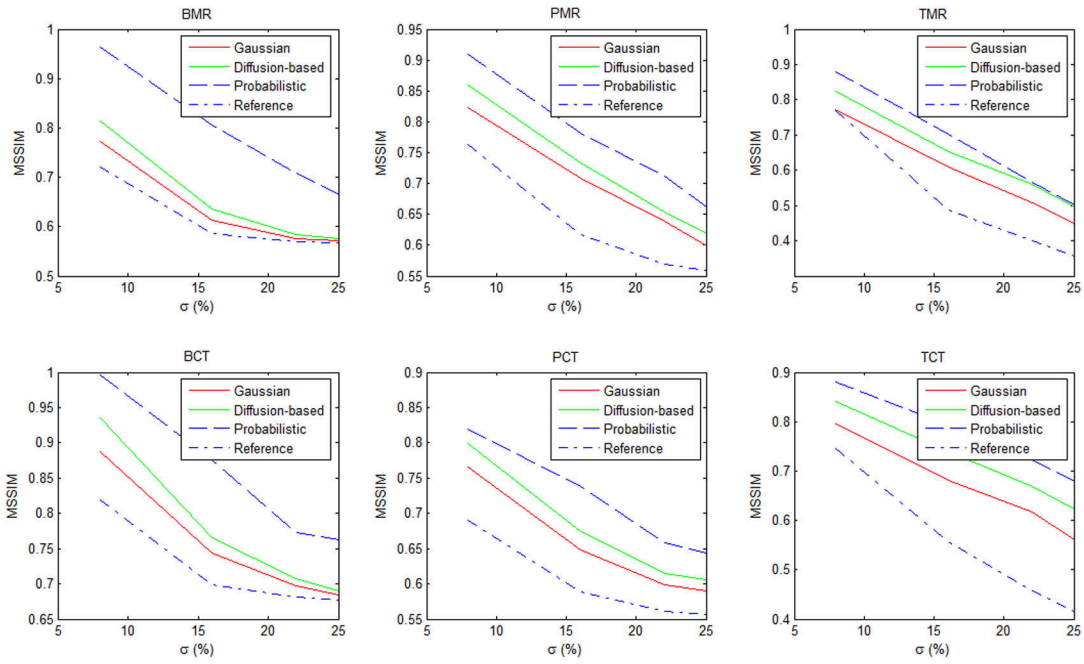

Figure 4.9: Plots of the MSSIM for the tested images (Fig. 3.11) under different noise scenarios. As with the PSNR results (Fig. 3.13), the MSSIM of the probabilistic complex phase representations is noticeably higher than the Gaussian and nonlinear diffusion-based representations.

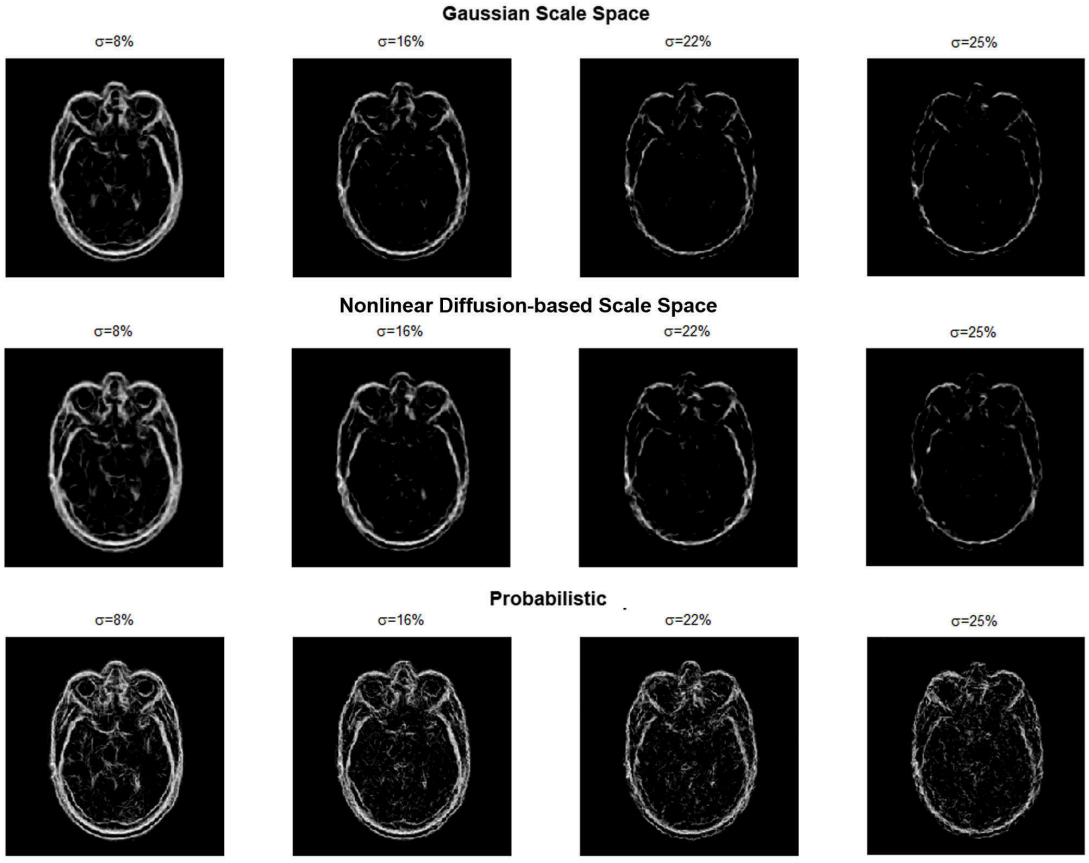

Figure 4.10: The multi-scale complex phase representations of the BMR image under the tested noise scenarios using Gaussian scale space, nonlinear diffusion-based scale space, and probabilistic framework.

Chapter 5

Probabilistic Complex Phase Representation Objective Function

This chapter describes in detail the overall design and performance analysis of the proposed probabilistic complex phase representation (PCPR) objective function. An introduction to statistical likelihood as an objective function in the context of comparing probabilistic complex phase representations is presented. An outlier sensitivity analysis is presented to examine different error distribution models estimators for use in the proposed PCPR objective function. The convergence smoothness of the proposed PCPR objective function is studied and examined in detail. Finally, the sensitivity of the proposed PCPR objective function to contrast non-uniformities and noise are studied.

5.1 Introduction

Given the probabilistic complex phase representation described in Chapter 4, the next step is to design an objective function that utilizes such representations for determining the alignment between images. There are several factors that need to be considered in the design of the objective function. First, the objective function should be robust to outliers such as noise and structural artifacts that can affect registration accuracy. Second, the objective function should ideally result in a monotonic cost function. This is important as local optimization schemes are dependent on the monotonicity of the underlying cost function to avoid getting trapped in local maxima. Finally, the objective function should be computationally efficient to allow for the registration of large images or large volumes of images within a reasonable time-frame.

5.2 Statistical Likelihood

The multimodal registration problem can be formulated as a maximization problem,

$$T_{\text{opt}} = \text{argmax}_T \left[\Psi \left(g(\underline{x}), f(T(\underline{x})) \right) \right], \tag{5.1}$$

where T_{opt} is the optimal transformation that aligns f and g, and $\Psi(.)$ is the objective function. To incorporate the concept of probabilistic complex phase representations into the multimodal registration problem, Eq. (5.1) can be rewritten as,

$$T_{\text{opt}} = \text{argmax}_T \left[\Psi \left(R_g(\underline{x}; t), R_f(T(\underline{x}); t) \right) \right]. \tag{5.2}$$

One effective approach for solving the registration problem in Eq. (5.2) is to perform maximum likelihood estimation, where the statistical likelihood of alignment given transformation T is employed as the objective function Ψ. Let us define the dissimilarity between the probabilistic complex phase representations at a point $\underline{x}$ and scale t as the residual error $\varepsilon(\underline{x}; t)$ between $R_g(\underline{x}; t)$ and $R_f(\underline{x}; t)$,

$$\varepsilon(\underline{x}; t) = R_g(\underline{x}; t) - R_f(\underline{x}; t). \tag{5.3}$$

Let the residual error $\varepsilon(\underline{x}; t)$ be considered a random variable following a probability distribution p given T,

$$\varepsilon(\underline{x}; t) \sim p(\varepsilon(\underline{x}; t); T). \tag{5.4}$$

The maximum likelihood estimate of T_{opt} can be found by maximizing the likelihood function $\mathcal{L}(T)$ over T,

$$\hat{T_{\text{opt}}} = \text{argmax}_T \, \mathcal{L}(T), \tag{5.5}$$

where the likelihood function $\mathcal{L}(T)$ is defined as

$$\mathcal{L}(T) = \prod_{\underline{x}} p(\varepsilon(\underline{x}; t); T). \tag{5.6}$$

The expression in Eq. (5.5) can equivalently be expressed as the minimization of the negative log likelihood,

$$\hat{T_{\text{opt}}} = \text{argmin}_T \left(- \sum_{\underline{x}} \log \left(p(\varepsilon(\underline{x}; t); T) \right) \right). \tag{5.7}$$

Based on Eq. (5.7), Huber [83] proposed that maximum likelihood estimation can be generalized further by replacing the negative log likelihood function with an estimation function ρ to form a generalized maximum likelihood-type estimate of T,

$$\hat{T_{\text{opt}}} = \text{argmin}_T \left(\sum_{\underline{x}} \rho(\varepsilon(\underline{x}; t); T) \right). \tag{5.8}$$

As such, the standard maximum likelihood estimation scheme described in Eq. (5.7) is a special case of the generalized maximum likelihood-type estimation scheme where $\rho(\varepsilon(\underline{x}; t); T) = -\log\left(p(\varepsilon(\underline{x}; t); T)\right)$.

The performance of the generalized maximum likelihood-type estimation scheme described in Eq. (5.8) depends heavily on the choice of ρ and can become highly biased when the residual error ε does not follow the probability distribution assumed by ρ. This situation is further complicated by the presence of outliers such as noise and structural artifacts that do not follow the error distribution assumed by ρ. As such, the estimation function ρ should be chosen such that the generalized maximum likelihood-type estimation scheme performs well in situations where the residual error ε follows the assumed distribution model but is not heavily influenced by situations where ε falls out of the assumed distribution model.

5.3 Error Distribution Model Analysis

As described in Section 5.2, the choice of the estimation function ρ is important to the overall performance of the generalized maximum likelihood-type estimation scheme, as it influences the assumed error distribution model. In the context of registration, the estimation function ρ should be chosen such that good discrimination is achieved for situations where the residual error follows the error distribution of the mutual structural characteristics between the images. Furthermore, the estimation ρ should be robust to outliers caused by noise and structural artifacts that fall outside of the mutual structural characteristics between the images. As such, a study of estimation functions in the context of outlier influence is important in choosing ρ, as well as indirectly defining the assumed error distribution model.

A commonly used estimation function in the literature is the quadratic estimation function,

$$\rho\left(\varepsilon(\underline{x};t);T\right) = \frac{\varepsilon(\underline{x};t)^2}{2\sigma^2}, \tag{5.9}$$

where σ is a scaling parameter. A special case of the quadratic estimation function described in Eq. (5.9) is the least squares estimation function when $\sigma = 1$. The quadratic estimation function is the optimal maximum likelihood estimate when the residual error ε follows a Gaussian distribution,

$$\begin{aligned} -\sum_{\underline{x}} \log\left(p(\varepsilon(\underline{x};t);T)\right) &= \sum_{\underline{x}} \rho\left(\varepsilon(\underline{x};t);T\right) \\ \log\left(p(\varepsilon(\underline{x};t);T)\right) &= -\rho\left(\varepsilon(\underline{x};t);T\right) \\ p(\varepsilon(\underline{x};t);T) &= \exp\left[-\rho\left(\varepsilon(\underline{x};t);T\right)\right] \\ p(\varepsilon(\underline{x};t);T) &= \exp\left[-\tfrac{\varepsilon(\underline{x};t)^2}{2\sigma^2}\right], \text{ if } \rho\left(\varepsilon(\underline{x};t);T\right) = \tfrac{\varepsilon(\underline{x};t)^2}{2\sigma^2}. \end{aligned} \tag{5.10}$$

There are two main advantages to using the quadratic estimation function. First, the quadratic estimation function provides good error discrimination for situations where the residual error follows the error distribution. Second, the quadratic estimation function is computationally efficient. The main problem with the quadratic estimation function is that it is very sensitive to outliers that deviate from the assumed error distribution model. One effective approach to studying the influence of outliers on an estimation function is study the behavior of the first derivative of the estimation function ρ with respect to the residual error ε [84]. The first derivative $\partial\rho/\partial\varepsilon$ of the quadratic estimation function ρ is,

$$\frac{\partial\rho}{\partial\varepsilon} = \frac{\varepsilon(\underline{x};t)}{\sigma^2}. \tag{5.11}$$

A plot of ρ (Eq. (5.9)) and the associated $\partial\rho/\partial\varepsilon$ (Eq. (5.11)) for the quadratic estimation function is shown in Fig. 5.1. The influence of outliers on the quadratic estimation function increases linearly and without bound [84]. This makes quadratic estimation functions behave poorly as the error distribution becomes heavy-tailed due to outliers such as noise and structural artifacts.

To increase robustness to outliers such as noise and structural artifacts while maintaining the good error discrimination of the quadratic estimation function, one approach is to design estimation functions whose derivatives, and hence the influence of outliers, tend to zero as $\varepsilon \to \pm\infty$. The main advantage of such redescending estimation functions [84–86] is that they are highly robust to large outliers that do not follow the assumed error distribution model while still taking outliers that are close to the assumed error distribution model

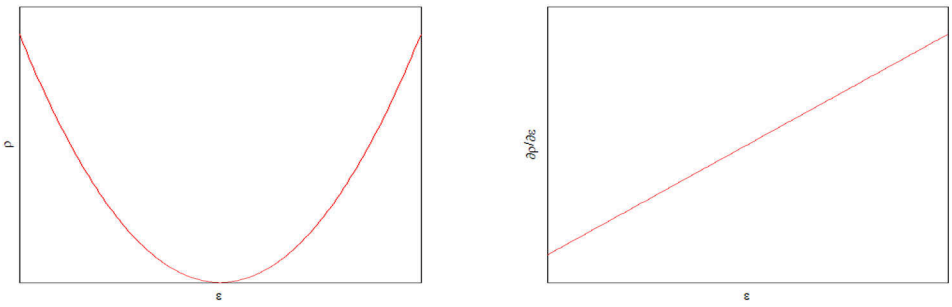

Figure 5.1: ρ and the associated $\partial\rho/\partial\varepsilon$ for the quadratic estimation function. The influence of outliers on the quadratic estimation function increases linearly and without bound [84]. This makes quadratic estimation functions behave poorly as the error distribution becomes heavy-tailed due to outliers such as noise and structural artifacts.

into account. To select an appropriate estimation function ρ for use in the generalized maximum likelihood-type estimation scheme to estimate T_{opt}, the influence of outliers is analyzed for the following estimation functions:

Tukey's Biweight [87] :

$$\rho(\varepsilon(\underline{x};t);T) = \begin{cases} \frac{\varepsilon(\underline{x};t)^2}{\sigma^2} - \frac{\varepsilon(\underline{x};t)^4}{\sigma^4} + \frac{\varepsilon(\underline{x};t)^6}{3\varepsilon(\underline{x};t)^6} & |\varepsilon(\underline{x};t)| \leq \sigma, \\ \frac{1}{3} & |\varepsilon(\underline{x};t)| > \sigma \end{cases} \tag{5.12}$$

$$\frac{\partial\rho}{\partial\varepsilon} = \begin{cases} \varepsilon(\underline{x};t)\left(1 - \frac{\varepsilon(\underline{x};t)^2}{\sigma^2}\right)^2 & |\varepsilon(\underline{x};t)| \leq \sigma, \\ 0 & |\varepsilon(\underline{x};t)| > \sigma \end{cases} \tag{5.13}$$

Hebert-Leahy [88] :

$$\rho(\varepsilon(\underline{x};t);T) = \log\left(1 + \frac{\varepsilon(\underline{x};t)^2}{\sigma^2}\right) \tag{5.14}$$

$$\frac{\partial\rho}{\partial\varepsilon} = \frac{\varepsilon(\underline{x};t)}{1 + \frac{\varepsilon(\underline{x};t)^2}{\sigma^2}} \tag{5.15}$$

Geman-McClure [89] :

$$\rho(\varepsilon(\underline{x};t);T) = \frac{\varepsilon(\underline{x};t)^2}{\sigma^2 + \varepsilon(\underline{x};t)^2} \tag{5.16}$$

$$\frac{\partial \rho}{\partial \varepsilon} = \frac{2\varepsilon(\underline{x};t)}{\sigma^2 + \varepsilon(\underline{x};t)^2} - \frac{2\varepsilon(\underline{x};t)^3}{(\sigma^2 + \varepsilon(\underline{x};t)^2)^2} \tag{5.17}$$

A plot of ρ and the associated $\partial\rho/\partial\varepsilon$ for the Tukey's Biweight, Hebert-Leahy, and Geman-McClure estimation functions are shown in Fig. 5.2. The influence of outliers on all of the redescending estimation functions, as indicated by $\partial\rho/\partial\varepsilon$, decreases towards zero as $\varepsilon \to 0$, thus indicating robustness to outliers that fall outside of the assumed error distribution. Of the three estimation functions, the Geman-McClure estimation function provides the best error discrimination in the areas that do fall within the assumed error distribution model, as evident by the sharper change in ρ with respect to ε near $\varepsilon = 0$ when compared to the other two estimation functions. Furthermore, based on the plots of $\partial\rho/\partial\varepsilon$, the Geman-McClure estimation function provides a good compromise between the hard outlier influence threshold exhibited by the Tukey's Biweight estimation function and the slow decrease in outlier influence exhibited by Hebert-Leahy estimation function. In terms of computational complexity, both the Tukey's Biweight and Geman-McClure estimation functions have relatively low computational complexity when compared to the Hebert-Leahy estimation function. Based on these observations, the Geman-McClure estimation function is chosen for ρ given its good error discrimination, good robustness to large outliers while still taking moderate outliers into account, and low computational complexity.

Since one of the main disadvantages of the quadratic estimation function is that it behaves poorly as the error distribution becomes heavy-tailed due to outliers such as noise and structural artifacts, the assumed error distribution model of the Geman-McClure estimation function in the context of the maximum likelihood framework to see how outliers are handled should be investigated. The error distribution $p(e(\underline{x};t);T)$ assumed by the Geman-McClure estimation function can be derived as follows,

$$\begin{aligned} -\sum_{\underline{x}} \log\left(p(\varepsilon(\underline{x};t);T)\right) &= \sum_{\underline{x}} \tfrac{\varepsilon(\underline{x};t)^2}{\sigma^2+\varepsilon(\underline{x};t)^2}, \text{ if } \rho\left((\underline{x};t);T\right) = \tfrac{\varepsilon(\underline{x};t)^2}{\sigma^2+\varepsilon(\underline{x};t)^2} \\ \log\left(p(\varepsilon(\underline{x};t);T)\right) &= -\tfrac{\varepsilon(\underline{x};t)^2}{\sigma^2+\varepsilon(\underline{x};t)^2} \\ p(\varepsilon(\underline{x};t);T) &= \exp\left[-\tfrac{\varepsilon(\underline{x};t)^2}{\sigma^2+\varepsilon(\underline{x};t)^2}\right]. \end{aligned} \tag{5.18}$$

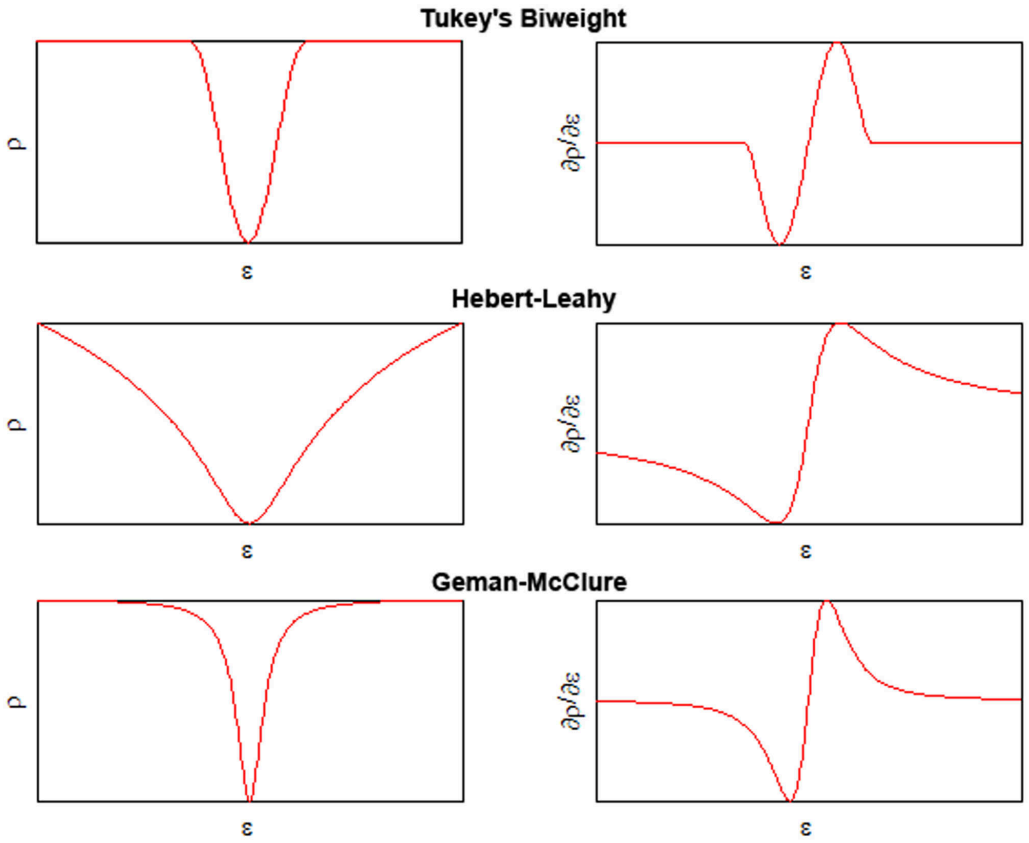

Figure 5.2: ρ and the associated $\partial\rho/\partial\varepsilon$ for the Tukey's Biweight [87], Herbert-Leahy [88], and Geman-McClure [89] estimation functions.

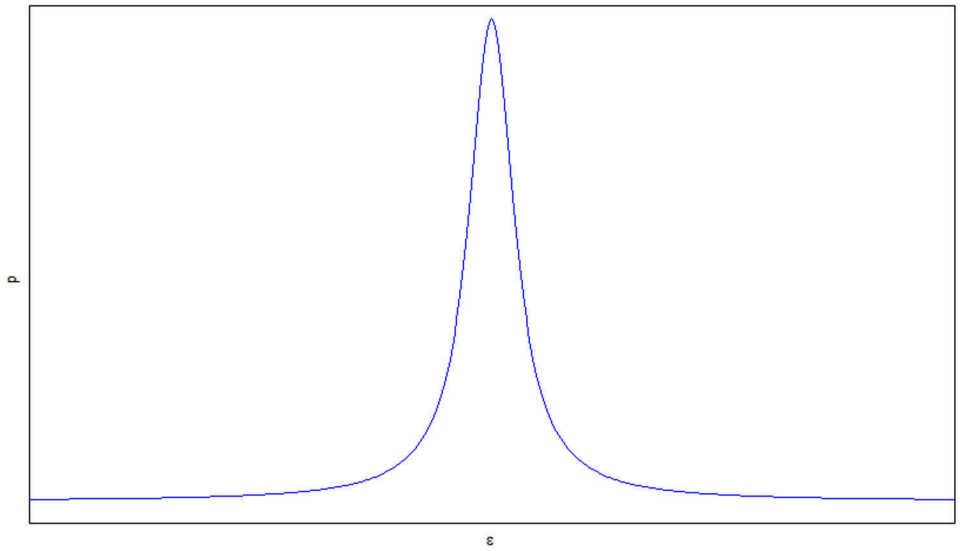

Figure 5.3: Error distribution assumed by the Geman-McClure estimation function as shown in Eq. (5.18). The error distribution has a similar shape to the Gaussian distribution in the regions near $\varepsilon = 0$. Therefore, it performs much like the quadratic estimation function when the residual error follows the Gaussian distribution model. However, the error distribution is characterized by a much heavier tail than the Gaussian distribution, thus taking the influence of outliers such as noise and structural artifacts into account in the error distribution model.

A plot of the error distribution $p(\varepsilon(\underline{x}; t); T)$ assumed by the Geman-McClure estimation function as shown in Eq. (5.18) is shown in Fig. 5.3. The error distribution has a similar shape to the Gaussian distribution in the regions near $\varepsilon = 0$. Therefore, it performs much like the quadratic estimation function when the residual error follows the Gaussian distribution model. However, the error distribution is characterized by a much heavier tail than the Gaussian distribution, thus taking the influence of outliers such as noise and structural artifacts into account in the error distribution model.

Given the Geman-McClure estimation function ρ within the generalized maximum likelihood-type estimation framework as well as taking the multi-scale nature of probabilistic representations into account, the proposed probabilistic complex phase representation (PCPR) objective function can be defined as the following likelihood function $\mathcal{L}(T)$ (Eq. (5.6)):

$$\Psi\left(g(\underline{x}), f(T(\underline{x}))\right) = \mathcal{L}(T) = \exp\left[-\left(\sum_{t=1}^{\gamma}\sum_{\underline{x}}\frac{\left(R_g(\underline{x};t) - R_f(T\left(\underline{x};\theta\right);t)\right)^2}{\sigma^2 + \left(R_g(\underline{x};t) - R_f(T\left(\underline{x};\theta\right);t)\right)^2}\right)\right]. \tag{5.19}$$

5.4 Convergence Smoothness

An important consideration in the design of an objective function is that the resulting monotonic cost function should ideally be monotonic to allow for local optimization schemes to smoothly converge to the global optima. As such, it is of great importance to study the cost functions produced by the proposed PCPR objective function based on probabilistic complex phase representations under different geometric distortions.

To study the cost function produces by the proposed PCPR objective function described in Eq. (5.19), a set of real medical image images acquired from the NLM Visible Human Project were translated horizontally within the ranges of $[-100, 100]$, $[-10, 10]$, and $[40, 50]$ (on a pixel basis) and, for a separate set of tests, rotated within the ranges of $[-50^o, 50^o]$, and $[-10^o, 10^o]$. The cost functions pertaining to the individual translation and rotation tests are then computed within these ranges. An overview of the translation and rotation tests are illustrated in Fig. 5.4 using a pair of example cost functions. A summary of the test image pairs is given below.

1. **BMC**: Brain, axial, 1mm resolution, PD-weighted MR and CT.
2. **PMC**: Pelvis, axial, 1mm resolution, T2-weighted MR and CT.
3. **TMC**: Torso, coronal, 1.875mm resolution, T1-weighted MR and CT.
4. **BMM**: Brain, axial, 1mm resolution, T1-weighted MR and T2-weighted MR.
5. **PMM**: Pelvis, coronal, 1mm resolution, T1-weighted MR and PD-weighted MR.

The test images are shown in Fig. 5.5. The normalized mutual information (NMI) [27] and generalized correlation ratio (CR) [34], both of which are considered to be state-of-the-art objective functions, were also tested for comparison purposes. Note that the cost functions produced by NMI and PCPR are negated such that the cost decreases as similarity between

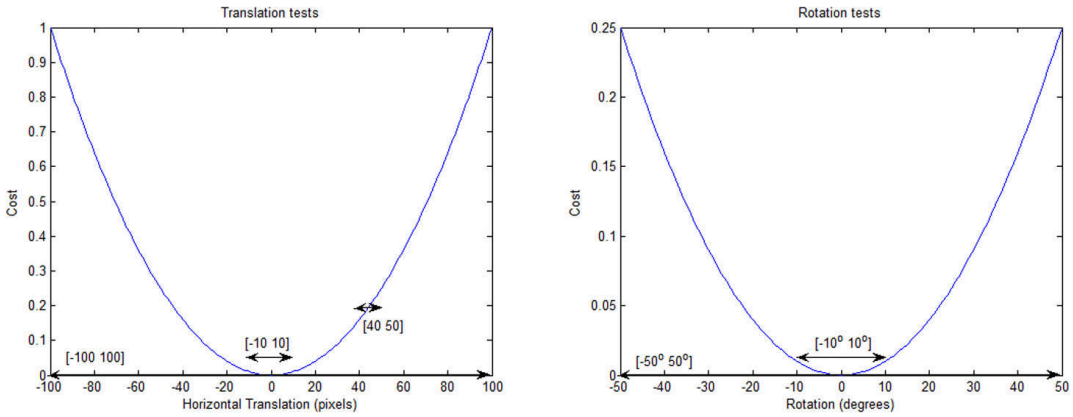

Figure 5.4: An overview of the translation and rotation tests are illustrated in Fig. 5.4 using a pair of example cost functions. The y-axis indicates the cost and the x-axis indicates the geometric distortion. The ground-truth alignment corresponds to zero translation and zero rotation respectively. Ideally, the cost function should be smooth and reach a global minimum at the zero point. The regions indicated by the double arrows indicate the individual translation and rotation tests.

the images increases. NMI was implemented using smoothed histograms computed with 100 intensity bins as specified by Mellor and Brady [41], while CR was implemented as specified by Roche et al. [34]. Trilinear interpolation was used in all experiments.

The cost functions for the translation ranges of $[-100, 100]$, $[-10, 10]$, and $[40, 50]$ using NMI [27], CR [34], and the proposed PCPR objective function are shown in Fig. 5.6, Fig. 5.7, and Fig. 5.8 respectively. The cost functions produced by all three objective functions are largely smooth and converge to the global optima, as illustrated in Fig. 5.6. These results demonstrate that the proposed PCPR objective function can provide monotonic cost functions to allow for local optimization schemes to smoothly converge to the global optima. The rate of change for the cost functions produced by the MI and PCPR objective functions remain high near the global optima for all cases, while the cost function produced by the CR objective function has a relatively low rate of change for the PMC, TMC, and BMM cases.

Fig. 5.7 and Fig. 5.8 provide a closer local view of the cost functions produced using the three objective functions. The cost functions produced by all three objective functions remain smooth and monotonic on the local scale, Fig. 5.8. A more telling story is told in Fig. 5.7, which shows the cost function near the global optima. The global optima

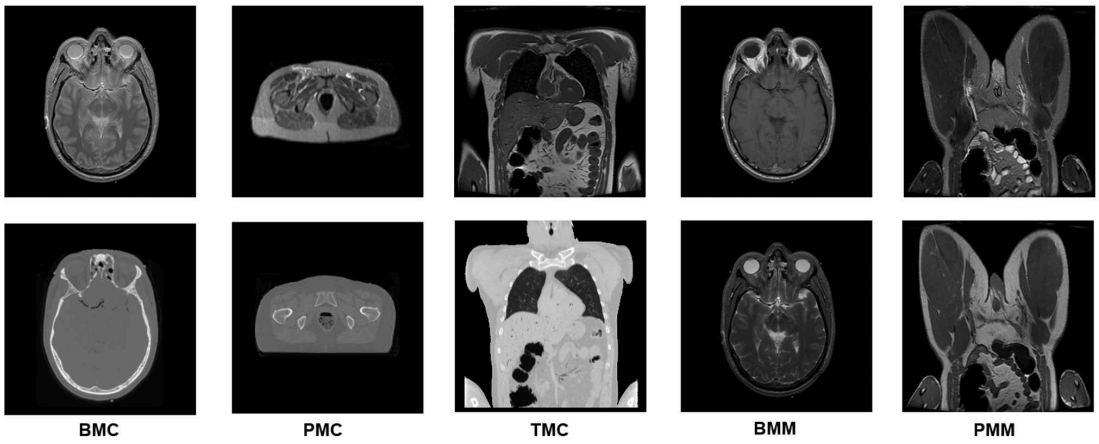

Figure 5.5: Test images used for convergence smoothness tests.

of the CR and PCPR objective functions coincide with the true alignment between the images for all cases. However, the global optima of the MI objective function does not correspond with the true alignment between the images for the BMC and PMC cases, which is undesirable for the purpose of image registration. These results demonstrate the effectiveness of the proposed PCPR objective function in determining the true alignment between images acquired using different imaging modalities.

The cost functions for the rotation ranges of $[-50^o, 50^o]$ and $[-10^o, 10^o]$ using NMI [27], CR [34], and the proposed PCPR objective function are shown in Fig. 5.9 and Fig. 5.10 respectively. The cost functions produced by the NMI and PCPR objective functions are largely smooth and converge to the global optima for all test cases, as shown in Fig. 5.9. Unfortunately, the cost functions produced by the CR objective function exhibit noticeable local optima in the TMC, BMM, and PMM cases. This indicates that the CR objective function can lead to local optimization schemes not converging to the global optima and is due to the fact that the functional mapping between the images used by the CR objective function can vary greatly as we converge to the global optima. Furthermore, as shown in Fig. 5.10, the global optima of the cost function produced by the CR objective function does not coincide with the true alignment for the BMC, PMC, and BMM cases. Fortunately, both the NMI and PCPR objective functions produced global optima that coincide with the true alignment of the images for all cases. These results demonstrate the effectiveness of the proposed PCPR objective function in determining the true alignment between images acquired using different imaging modalities.

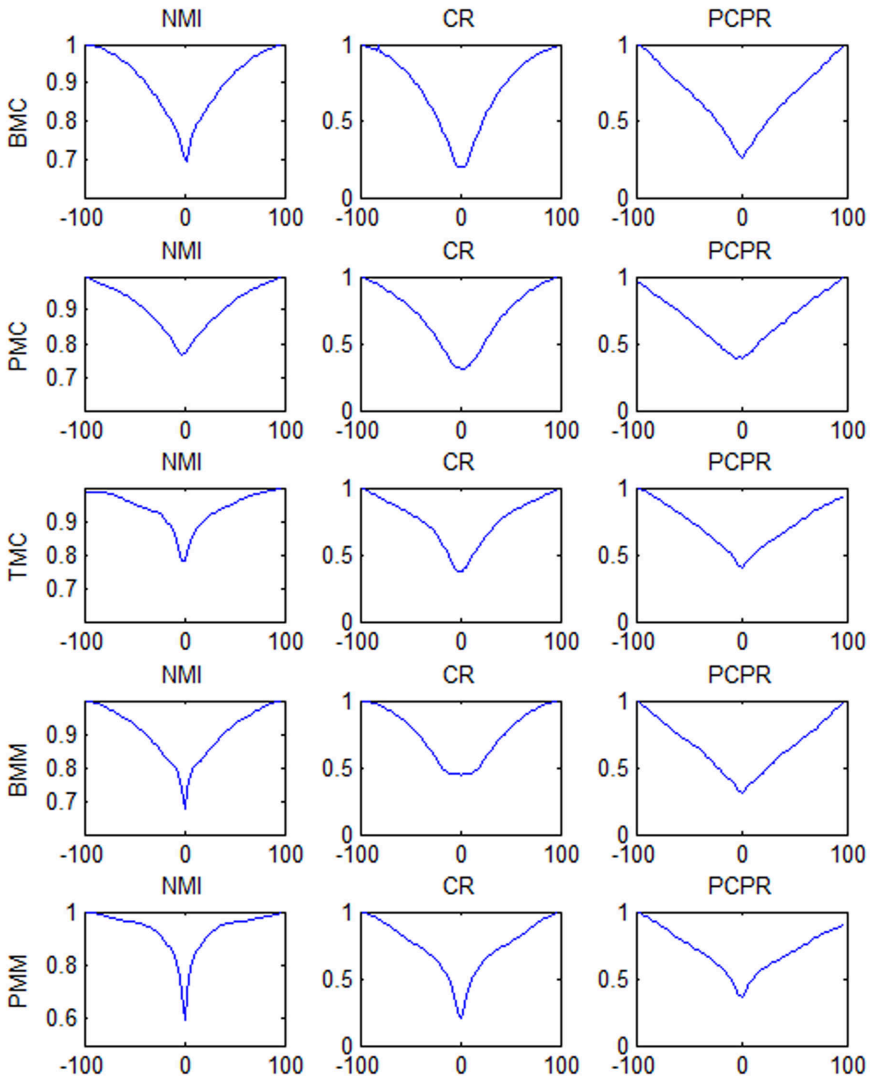

Figure 5.6: Cost functions for the range of -100 to 100 for the tested images using NMI [27], CR [34], and the proposed PCPR objective function. The cost functions produced by all three objective functions are largely smooth and converge to the global optima.

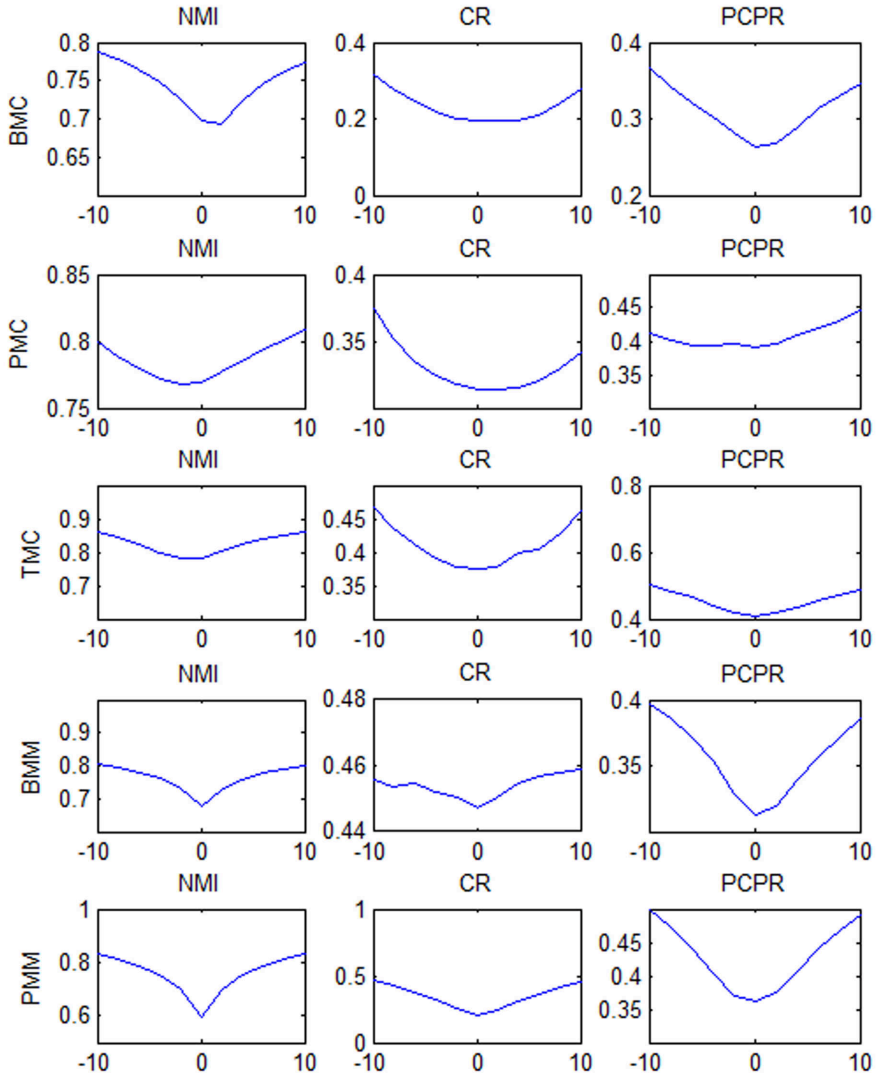

Figure 5.7: Cost functions for the range of -10 to 10 for the tested images using NMI [27], CR [34], and the proposed PCPR objective function. The global optima of the CR and PCPR objective functions coincide with the true alignment between the images for all cases. However, the global optima of the MI objective function does not correspond with the true alignment between the images for the BMC and PMC cases, which is undesirable for the purpose of image registration.

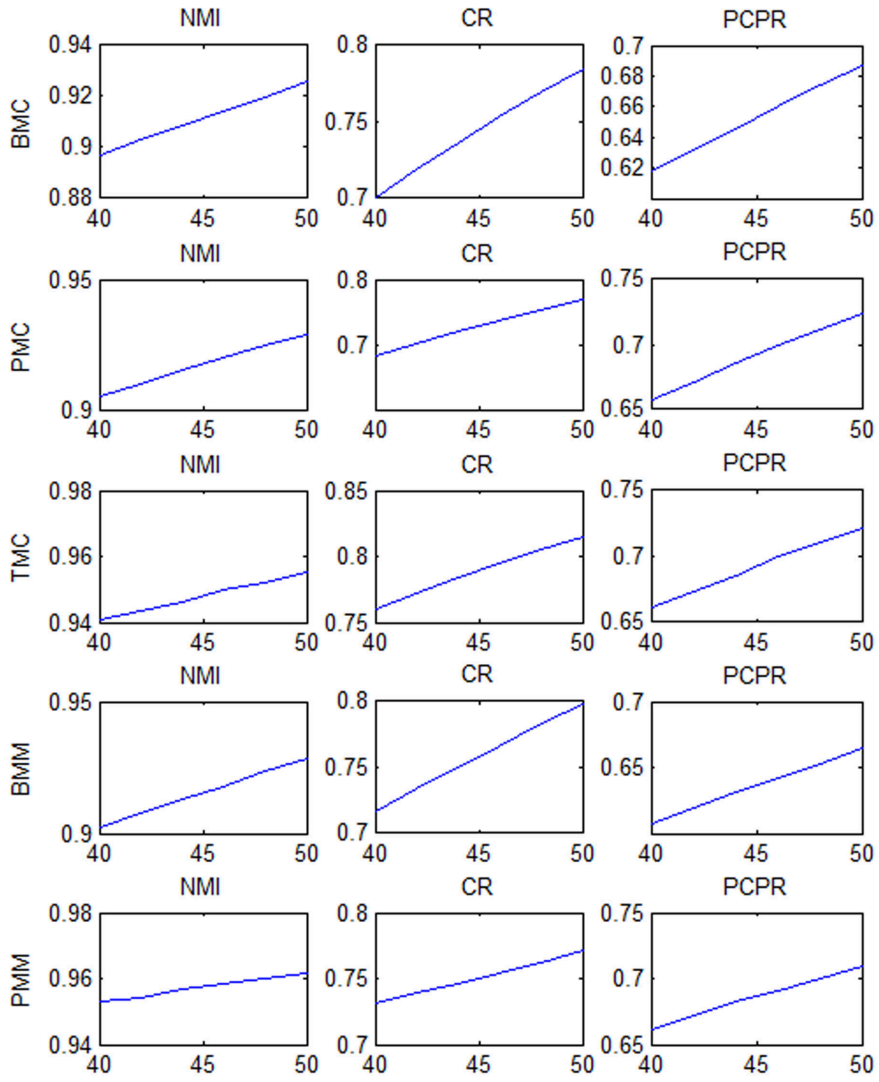

Figure 5.8: Cost functions for the range of 40 to 50 for the tested images using NMI [27], CR [34], and the proposed PCPR objective function. The cost functions produced by all three objective functions remain smooth and monotonic on the local scale.

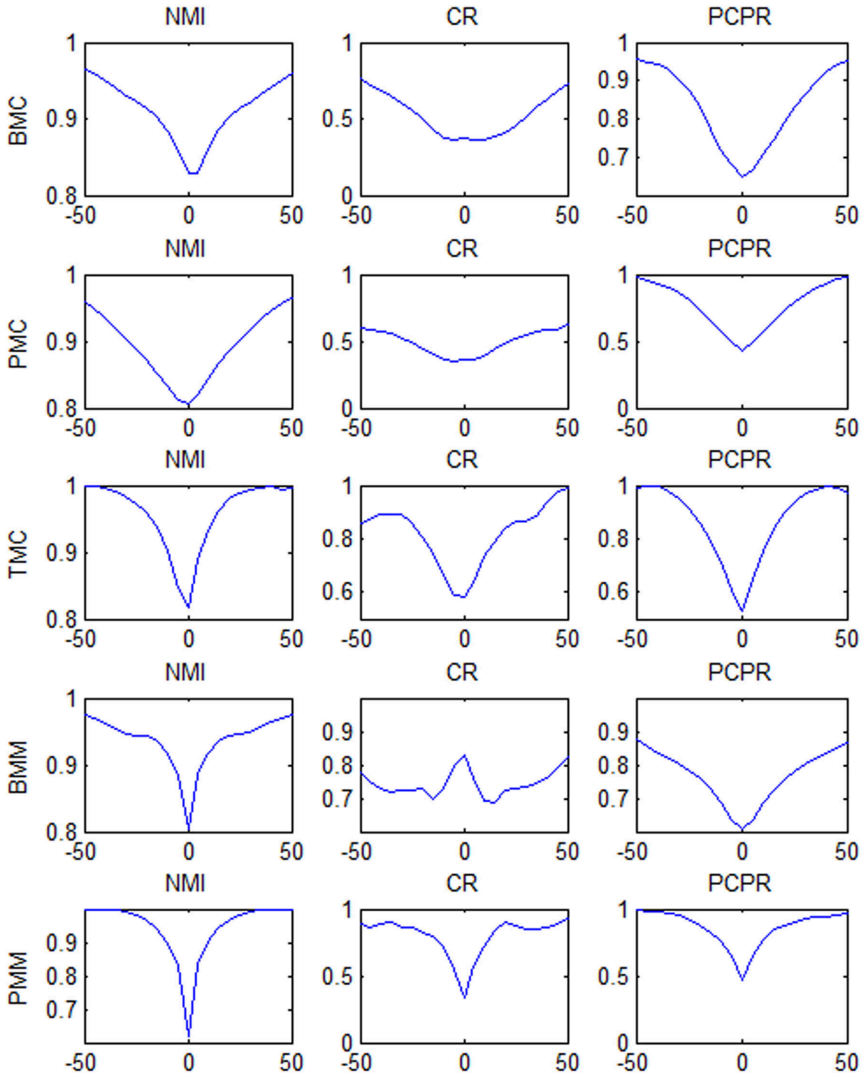

Figure 5.9: Cost functions for the range of -50^o to 50^o for the tested images using NMI [27], CR [34], and the proposed PCPR objective function. The cost functions produced by the NMI and PCPR objective functions are largely smooth and converge to the global optima for all test cases. Unfortunately, the cost functions produced by the CR objective function exhibit noticeable local optima in the TMC, BMM, and PMM cases.

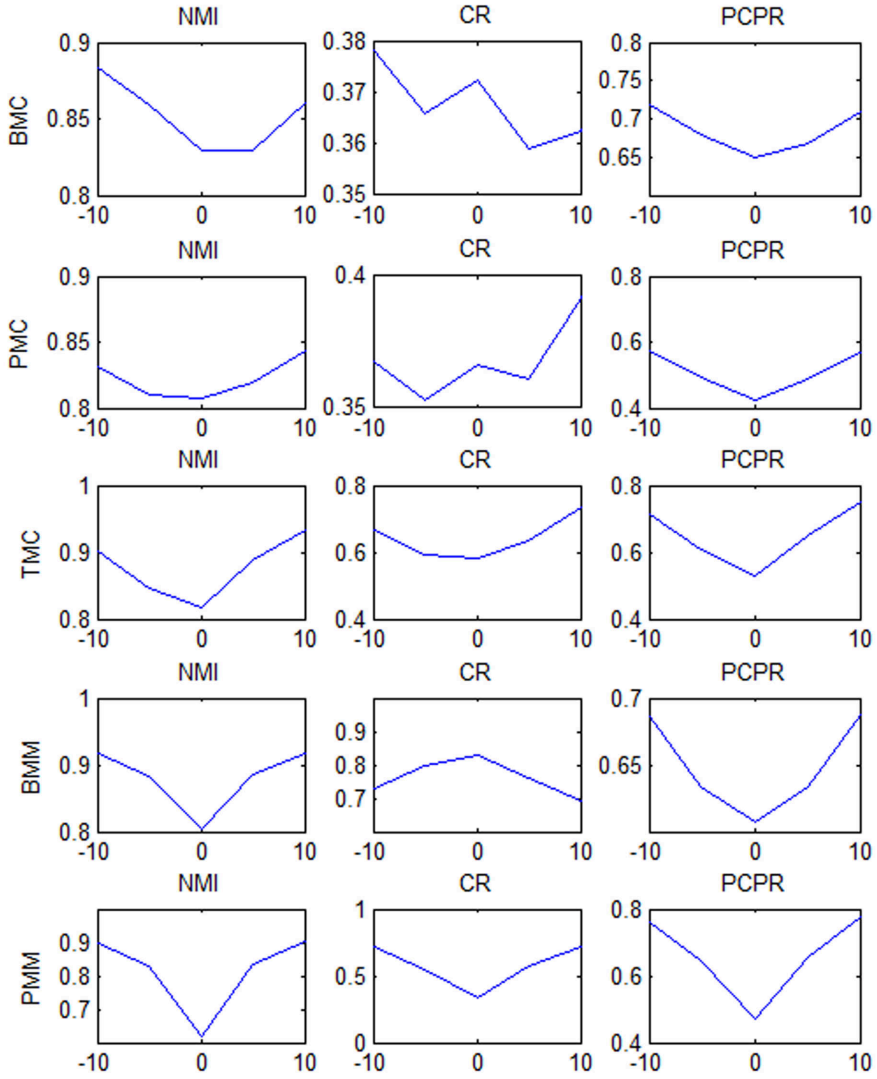

Figure 5.10: Cost functions for the range of -10^o to -10^o for the tested images using NMI [27], CR [34], and the proposed PCPR objective function. The global optima of the cost function produced by the CR objective function does not coincide with the true alignment for the BMC, PMC, and BMM cases. Fortunately, both the NMI and PCPR objective functions produced global optima that coincide with the true alignment of the images for all cases.

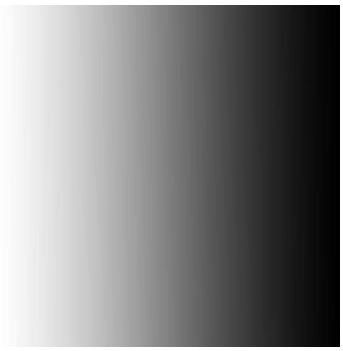

Figure 5.11: Gradient mask used in contrast non-uniformities test.

5.5 Non-uniformity Sensitivity

Another important consideration in the design of an objective function is robustness to contrast non-uniformities that can affect registration accuracy. To provide robust registration in the presence of contrast non-uniformities, the underlying objective function should be ideally remain smooth and monotonically decrease towards the global optima even under the influence of contrast uniformities. As such, it is important to study the cost functions produced by the proposed probabilistic complex phase representation (PCPR) objective function based on probabilistic complex phase representations under contrast non-uniformities.

To study the cost function produced by the proposed PCPR objective function described in Eq. (5.19) under different contrast non-uniformities, the set of tests conducted in Section 5.6 were performed after a gradient mask was applied to the images to induce contrast non-uniformities. The gradient mask used is shown in Fig. 5.11. The gradient mask is first applied to the first image in each test set and then flipped horizontally before being applied to the second image. The PMM test images after inducing contrast non-uniformities using the gradient mask is shown in Fig. 5.12.

The cost functions for the translation ranges of $[-100, 100]$, $[-10, 10]$, and $[40, 50]$ after contrast non-uniformities using NMI [27], CR [34], and the proposed PCPR objective function are shown in Fig. 5.13, Fig. 5.14, and Fig. 5.15 respectively. From Fig. 5.13, the cost functions produced by NMI is largely smooth and converge to the global optima in all but the PMM test case, where a local optimum is noticeable in the region corresponding to right translations. Similarly, the cost functions produced by PCPR is largely smooth and converge to the global optima in all but the PMM test case, where a local optimum is

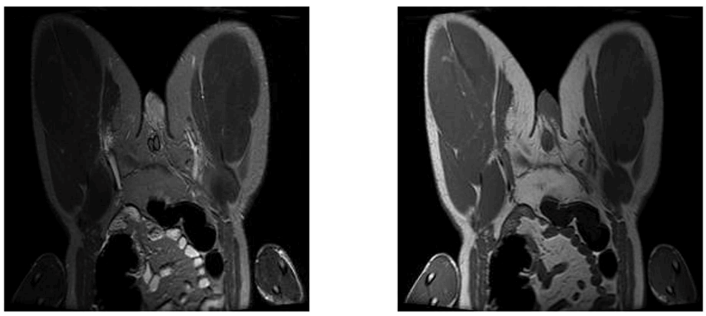

Figure 5.12: The PMM test images after inducing contrast non-uniformities using the gradient mask.

noticeable in the region corresponding to right translations, although the local optima is noticeably less steep than that produced by NMI. Furthermore, a large and steep optimal optima is seen in the cost functions produced by CR cost function for the TMC case, and the global optima is located far from the true alignment in the PMM case. This indicates that the CR objective function performs poorly when compared to NMI and PCPR in handling contrast non-uniformities and is due to the fact that the CR objective function relies on a functional intensity relationship between the images being registered, which may not exist due to the presence of contrast non-uniformities.

Fig. 5.14 and Fig. 5.15 provide a closer local view of the cost functions produced using the three objective functions under contrast non-uniformities. From Fig. 5.14, the global optima produced by NMI and PCPR coincide with the true alignment of the images. Unfortunately, from Fig. 5.15, the local optima exist in the cost function produced by NMI and PCPR for the PMM case, indicating that while both methods are robust to contrast non-uniformities, they are by no means perfect in terms of convergence smoothness under such situations.

The cost functions for the rotation ranges of $[-50^o, 50^o]$ and $[-10^o, 10^o]$ after contrast non-uniformities using NMI [27], CR [34], and the proposed PCPR objective function are shown in Fig. 5.16 and Fig. 5.17 respectively. The cost functions produced by the NMI and PCPR objective functions are largely smooth and converge to the global optima for all but the PMM test case, where there is a local optima at the region corresponding to right rotation. Unfortunately, the cost functions produced by the CR objective function exhibit noticeable local optima in the TMC, BMM, and PMM cases. Furthermore, as shown in

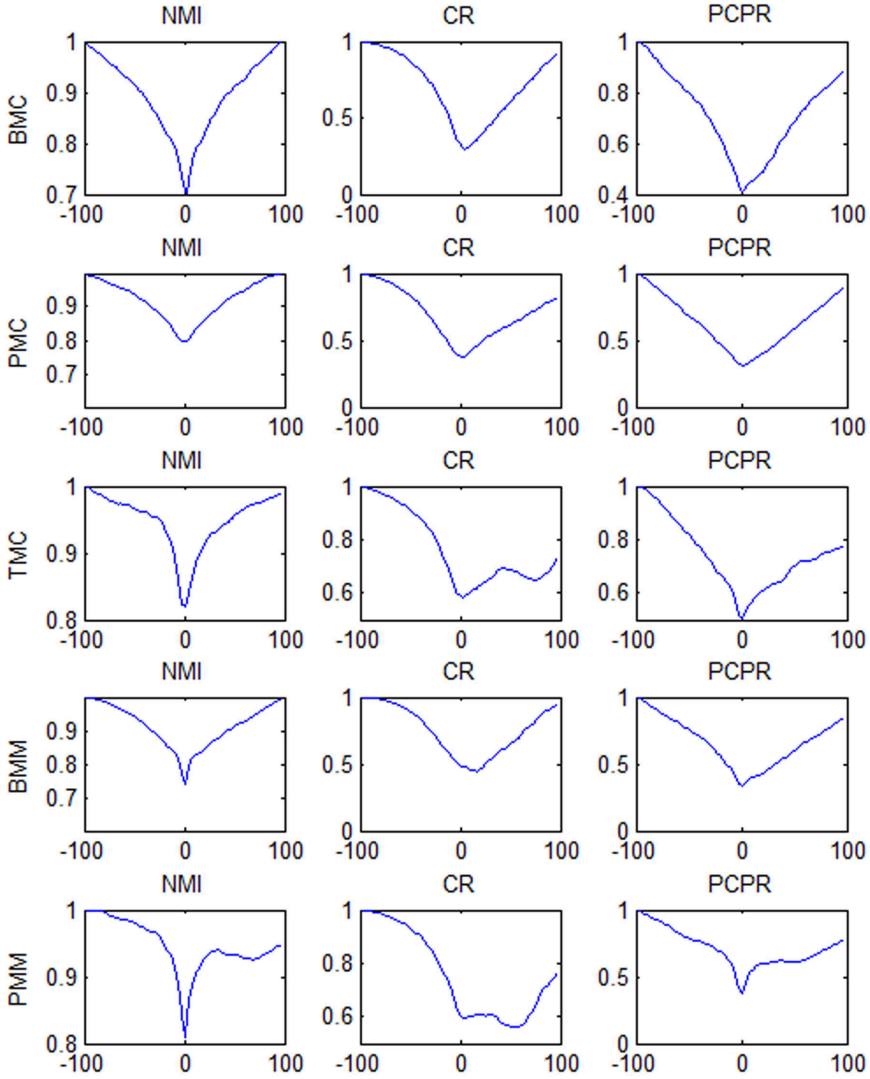

Figure 5.13: Cost functions for the range of -100 to 100 for the tested images after contrast non-uniformities using NMI [27], CR [34], and the proposed PCPR objective function. The cost functions produced by NMI and PCPR are largely smooth and converge to the global optima in all but the PMM test case, where a local optimum is noticeable in the region corresponding to right translations. Furthermore, a large and steep optimal optima is seen in the cost functions produced by CR cost function for the TMC case, and the global optima is located far from the true alignment in the PMM case.

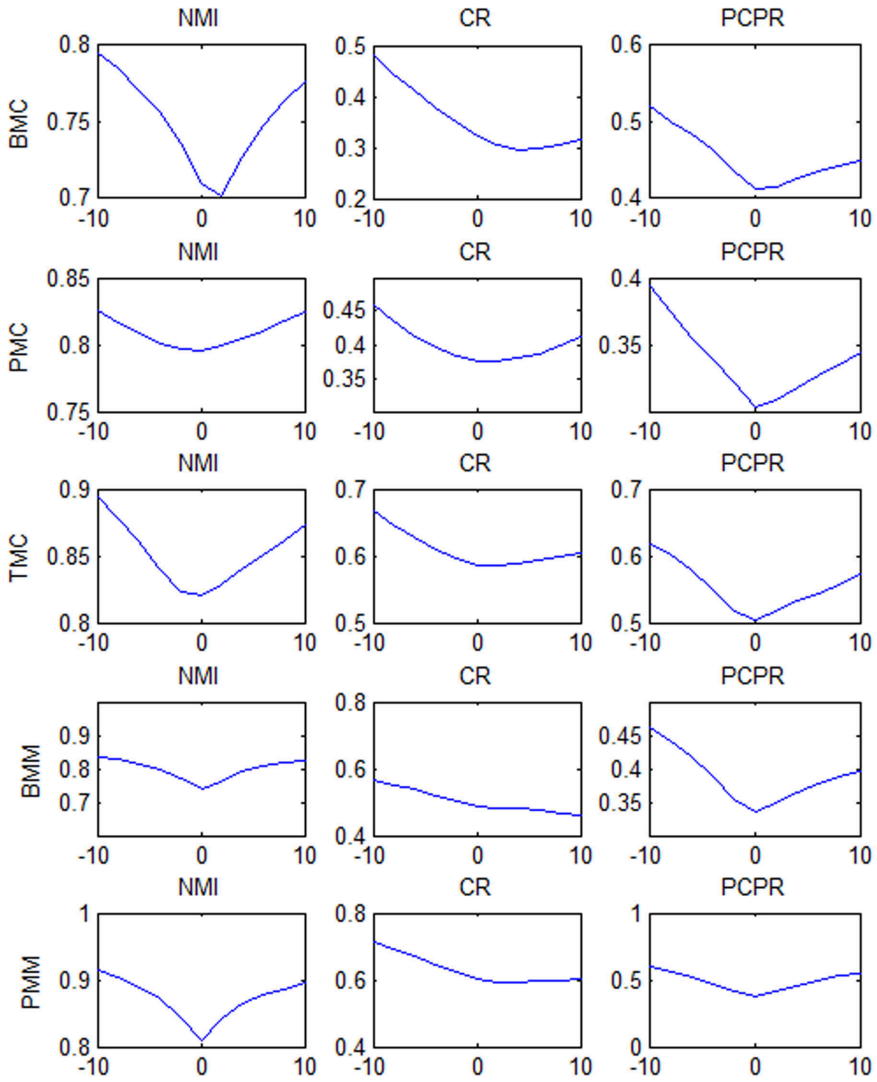

Figure 5.14: Cost functions for the range of -10 to 10 for the tested images after contrast non-uniformities using NMI [27], CR [34], and the proposed PCPR objective function. The global optima produced by NMI and PCPR coincide with the true alignment of the images.

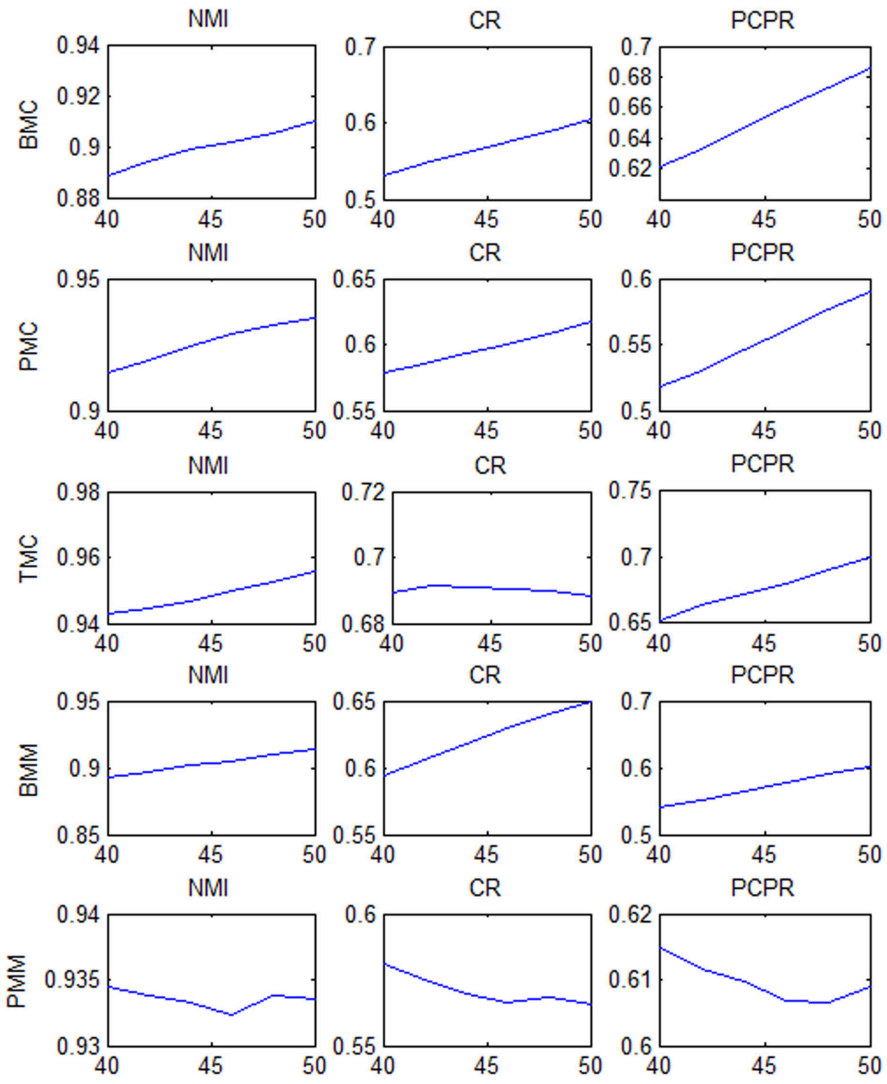

Figure 5.15: Cost functions for the range of 40 to 50 for the tested images after contrast non-uniformities using NMI [27], CR [34], and the proposed PCPR objective function. Unfortunately, the local optima exist in the cost function produced by NMI and PCPR for the PMM case.

Fig. 5.17, the global optima of the cost function produced by the CR objective function does not coincide with the true alignment for the BMC, PMC, and BMM cases. Fortunately, both the NMI and PCPR objective functions produced global optima that coincide with the true alignment of the images for all cases. These results demonstrate the effectiveness of the proposed PCPR objective function in determining the true alignment between images acquired using different imaging modalities under contrast non-uniformities.

5.6 Noise sensitivity

Another important consideration in the design of an objective function is robustness to noise artifacts that can affect registration accuracy. To provide robust registration in the presence of noise, the underlying objective function should be ideally remain smooth and monotonically decrease towards the global optima even under the influence of high noise levels. As such, it is of great importance to study the cost functions produced by the proposed probabilistic complex phase representation (PCPR) objective function based on probabilistic complex phase representations under different noise levels.

To study the cost function produced by the proposed PCPR objective function described in Eq. (5.19) under different noise levels, the set of tests conducted in Section 5.6 were performed under different additive Gaussian noise scenarios with standard deviations of $\sigma = \{8\%, 16\%, 22\%\}$.

The cost functions for the translation ranges of $[-100, 100]$, $[-10, 10]$, and $[40, 50]$ for $\sigma = 8\%$ are shown in Fig. 5.18, Fig. 5.19, and Fig. 5.20 respectively. From Fig. 5.18, the cost function produced by the NMI objective function begins to exhibit noticeable local optima in all cases, with the BMM case exhibiting global optima far away from the true alignment of the images. Furthermore, the presence of local optima in the cost functions produced using the NMI objective function is seen in Fig. 5.20. This is due to the fact that the joint and marginal intensity distributions computed in the presence of noise is not representative of the statistical characteristics of the underlying image content. Since the computation of NMI relies on reliable joint and marginal intensity distributions, the use of the NMI objective function is highly sensitive to the presence of noise. The cost functions produced by the CR and PCPR objective functions remain smooth and monotonically decreasing towards the global optima in most cases, with the exception of noticeable local optima in the TMC and BMM cases for the CR objective function seen in Fig. 5.18.

Fig. 5.19 shows the cost function near the global optima for $\sigma = 8\%$. The global optima of the CR and PCPR objective functions coincide with the true alignment between the

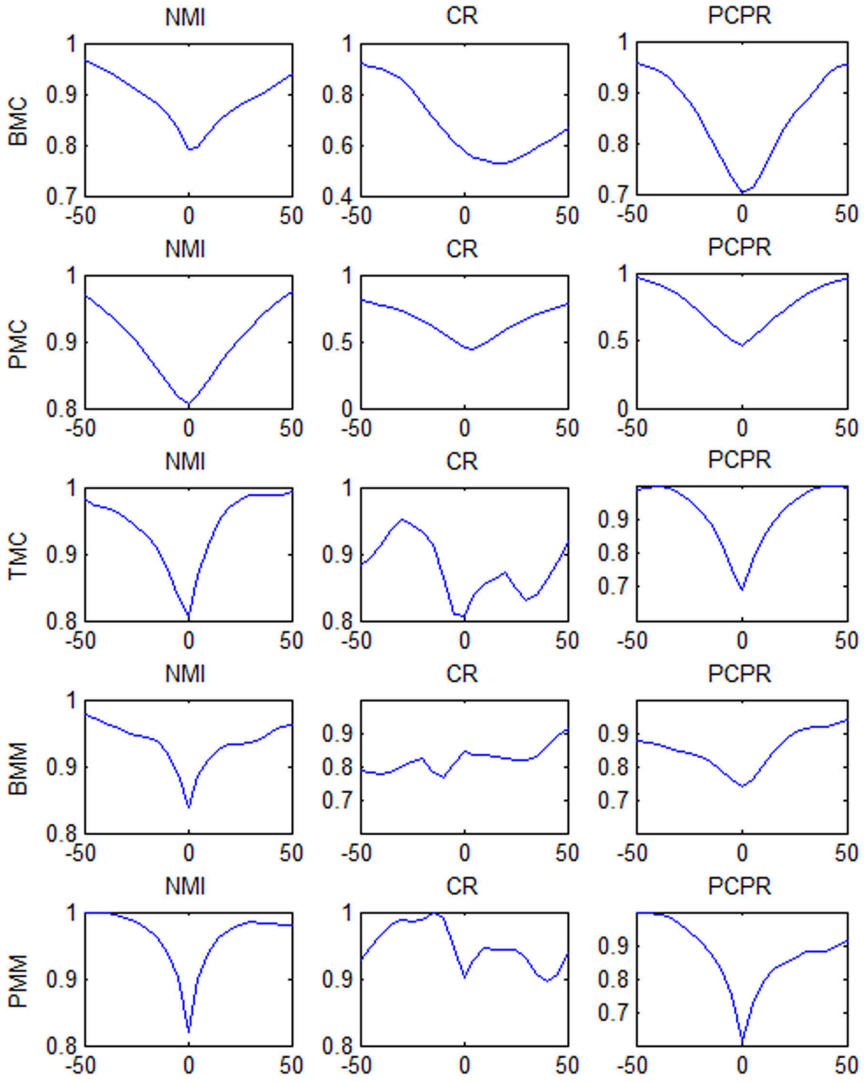

Figure 5.16: Cost functions for the range of -50^o to 50^o for the tested images after contrast non-uniformities using NMI [27], CR [34], and the proposed PCPR objective function. The cost functions produced by NMI and PCPR are largely smooth and converge to the global optima for all but the PMM test case, where there is a local optima at the region corresponding to right rotation. Unfortunately, the cost functions produced by CR exhibit noticeable local optima in the TMC, BMM, and PMM cases.

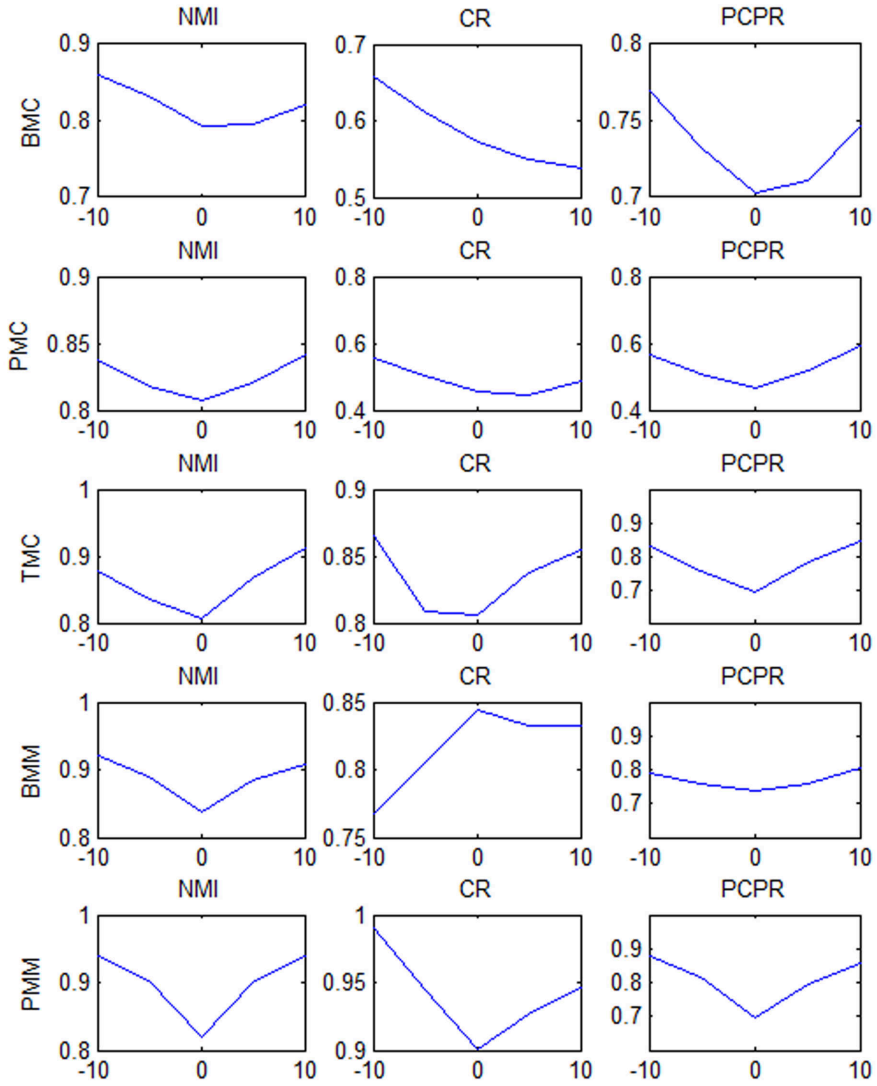

Figure 5.17: Cost functions for the range of -10^o to -10^o for the tested images after contrast non-uniformities using NMI [27], CR [34], and the proposed PCPR objective function. The global optima of the cost function produced by CR does not coincide with the true alignment for the BMC, PMC, and BMM cases. Fortunately, both NMI and PCPR produced global optima that coincide with the true alignment of the images for all cases.

images for all cases, with the exception of the PMC case for CR, showing their robustness to noise. However, the global optima of the cost functions produced by the NMI objective function no longer coincide with the true alignment of the images for the BMC, PMC, and TMC cases.

The cost functions for the translation ranges of $[-100, 100]$, $[-10, 10]$, and $[40, 50]$ for $\sigma = 16\%$ are shown in Fig. 5.21, Fig. 5.22, and Fig. 5.23 respectively. From Fig. 5.21, the cost functions produced by the NMI objective function exhibit a strange phenomena, where the cost function begins to indicate increasing dissimilarity between the images as they move into alignment. This is due to the fact that the statistical characteristics of noise has corrupted the statistical characteristics of the underlying image content to a point where the alignment of the images actually results in reduced mutual information. Furthermore, the cost functions produced by the NMI objective function exhibit noticeable local optima both at the global scale and at the local scale seen in Fig. 5.22 and Fig. 5.23. As with the case of $\sigma = 8\%$, the cost functions produced by the CR and PCPR objective functions are largely smooth and monotonically decreasing towards the global optima, with the exception of noticeable local optima in the BMM case for the CR objective function seen in Fig. 5.21 on the global scale and Fig. 5.22 on the local scale near the true alignment. These results demonstrate that the proposed PCPR objective function may be more robust to high levels of noise than the CR and NMI objective functions.

The cost functions for the translation ranges of $[-100, 100]$, $[-10, 10]$, and $[40, 50]$ for $\sigma = 22\%$ are shown in Fig. 5.24, Fig. 5.25, and Fig. 5.26 respectively. NMI performs very poorly under this high noise scenario, exhibiting noticeable local optima and increasing dissimilarity as the images come into alignment. The cost functions produced by CR now exhibit local optima near the true alignment for the BMC case, in addition to the local optima present in the BMM case. As with the previous noise levels, the cost function produced by PCPR remain largely smooth and converge to the true alignment of the images. These results further reinforce the claim of noise robustness of PCPR in the context of multimodal image registration.

5.7 Summary

In this chapter, an objective function based on probabilistic complex phase representations was presented. The proposed objective function was designed to address the issues associated with robustness to outliers such as noise and modality-related artifacts. A study of convergence smoothness demonstrated that the proposed objective function produced cost

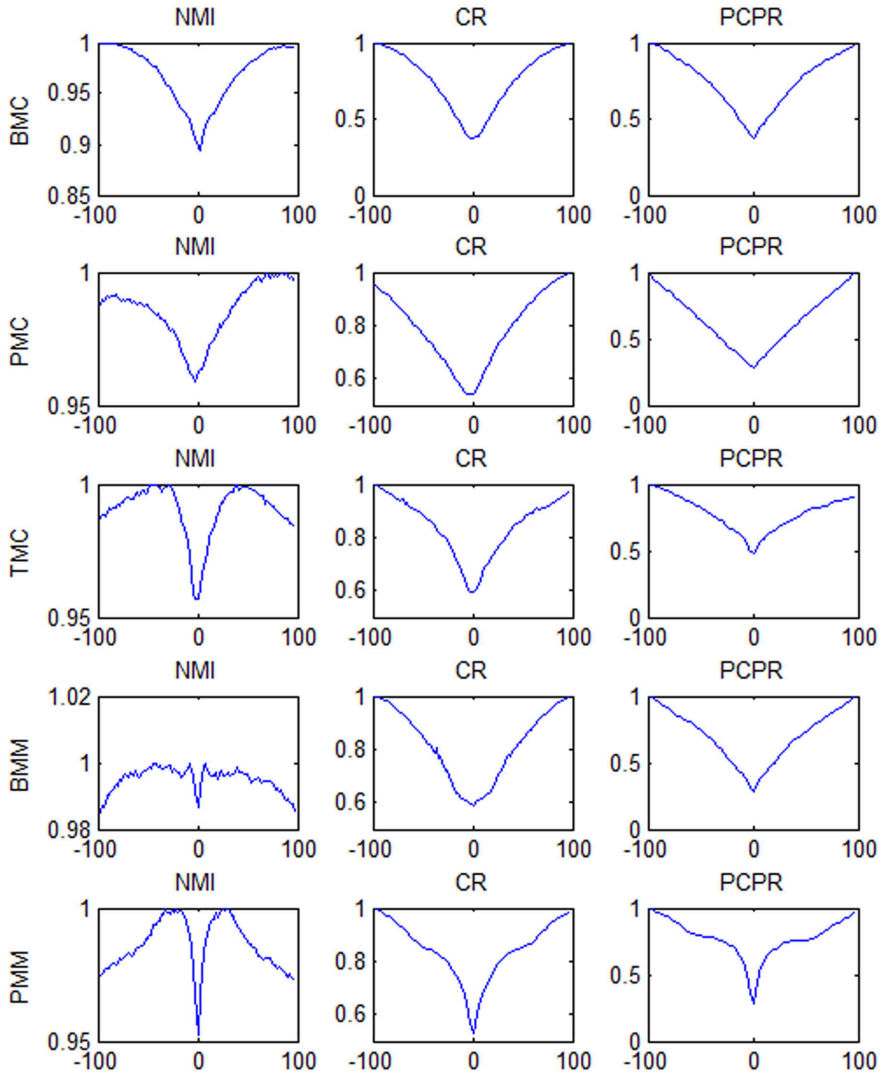

Figure 5.18: Cost functions for the range of -100 to 100 for $\sigma = 8\%$ for the tested images using NMI [27], CR [34], and the proposed PCPR objective function. The cost function produced by NMI begins to exhibit noticeable local optima in all cases, with the BMM case exhibiting global optima far away from the true alignment of the images. The cost functions produced by CR and PCPR remain smooth and monotonically decreasing towards the global optima in most cases, with the exception of noticeable local optima in the TMC and BMM cases for the CR objective function.

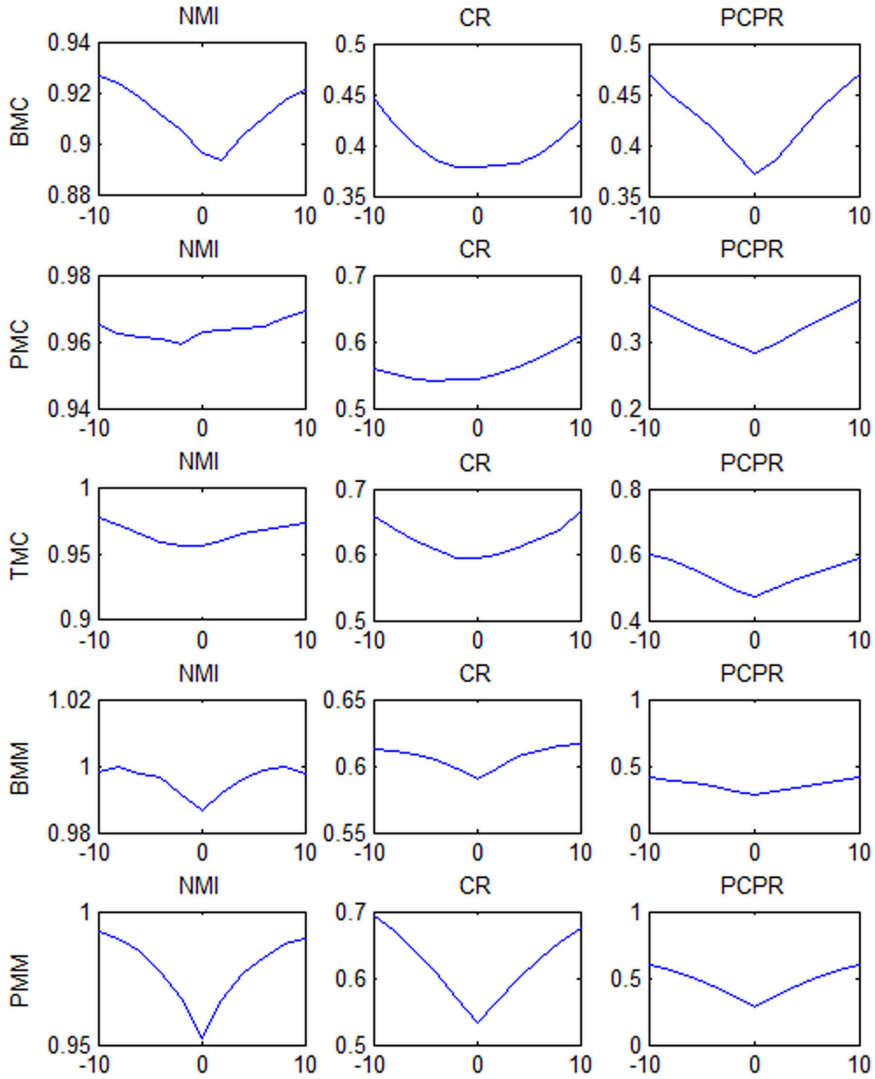

Figure 5.19: Cost functions for the range of -10 to 10 for $\sigma = 8\%$ for the tested images using NMI [27], CR [34], and the proposed PCPR objective function. The global optima of CR and PCPR coincide with the true alignment between the images for all cases, with the exception of the PMC case for CR, showing their robustness to noise. However, the global optima of the cost functions produced by NMI no longer coincide with the true alignment of the images for the BMC, PMC, and TMC cases.

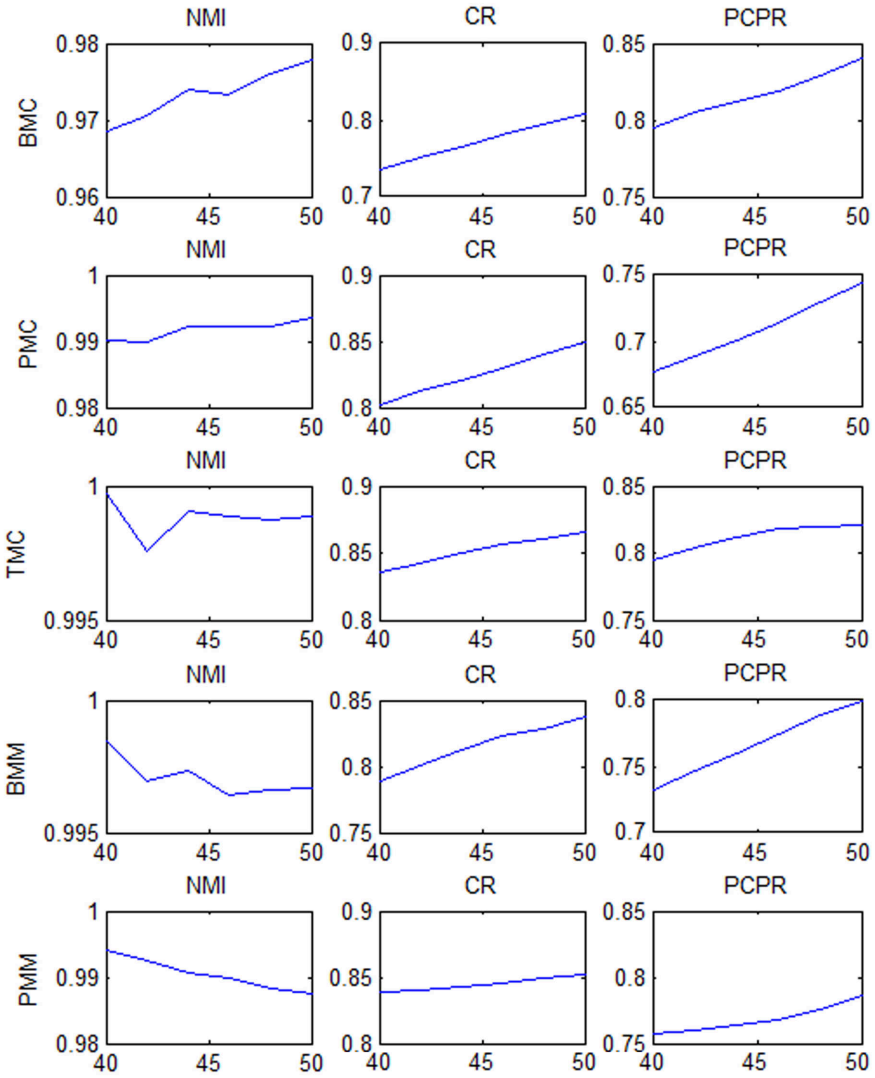

Figure 5.20: Cost functions for the range of 40 to 50 for $\sigma = 8\%$ for the tested images using NMI [27], CR [34], and the proposed PCPR objective function. Local optima is present in the cost functions produced using the NMI objective function.

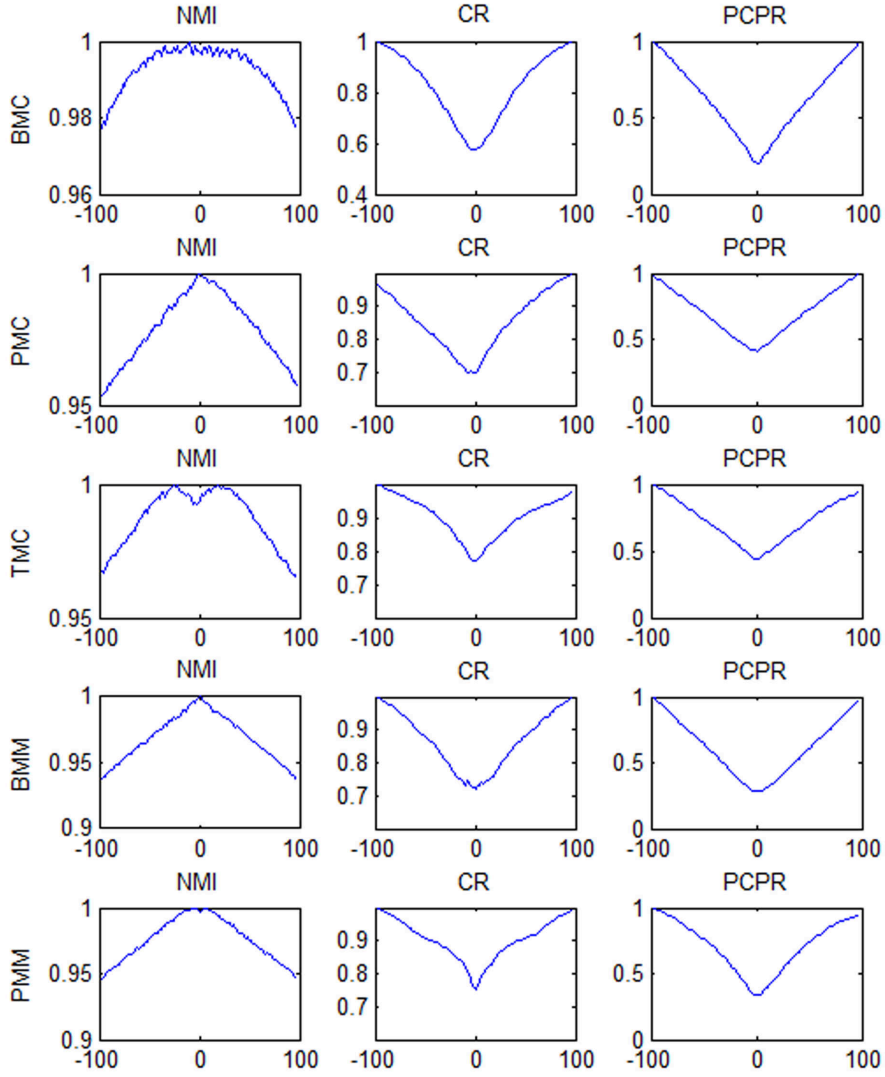

Figure 5.21: Cost functions for the range of -100 to 100 for $\sigma = 16\%$ for the tested images using NMI [27], CR [34], and the proposed PCPR objective function. The cost functions produced by NMI exhibit a strange phenomena, where the cost function begins to indicate increasing dissimilarity between the images as they move into alignment. The cost functions produced by CR and PCPRare largely smooth and monotonically decreasing towards the global optima, with the exception of noticeable local optima in the BMM case for the CR objective function.

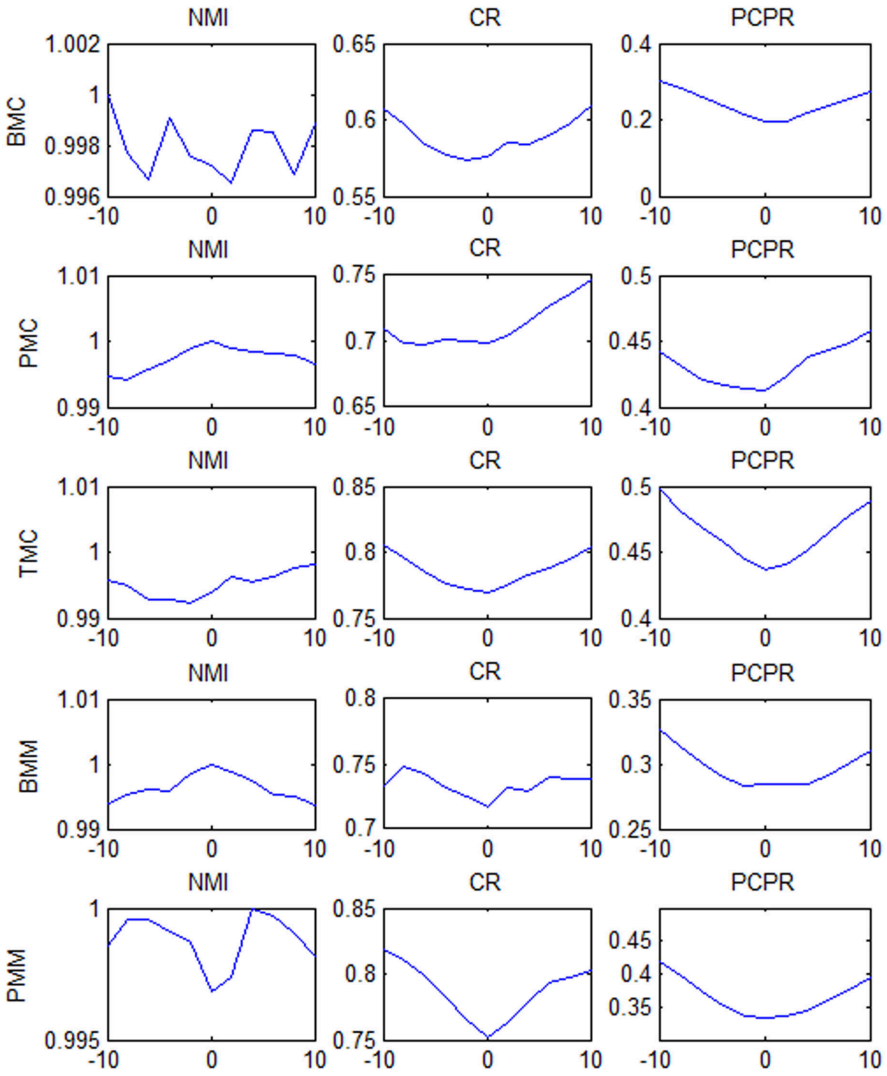

Figure 5.22: Cost functions for the range of -10 to 10 for $\sigma = 16\%$ for the tested images using NMI [27], CR [34], and the proposed PCPR objective function.

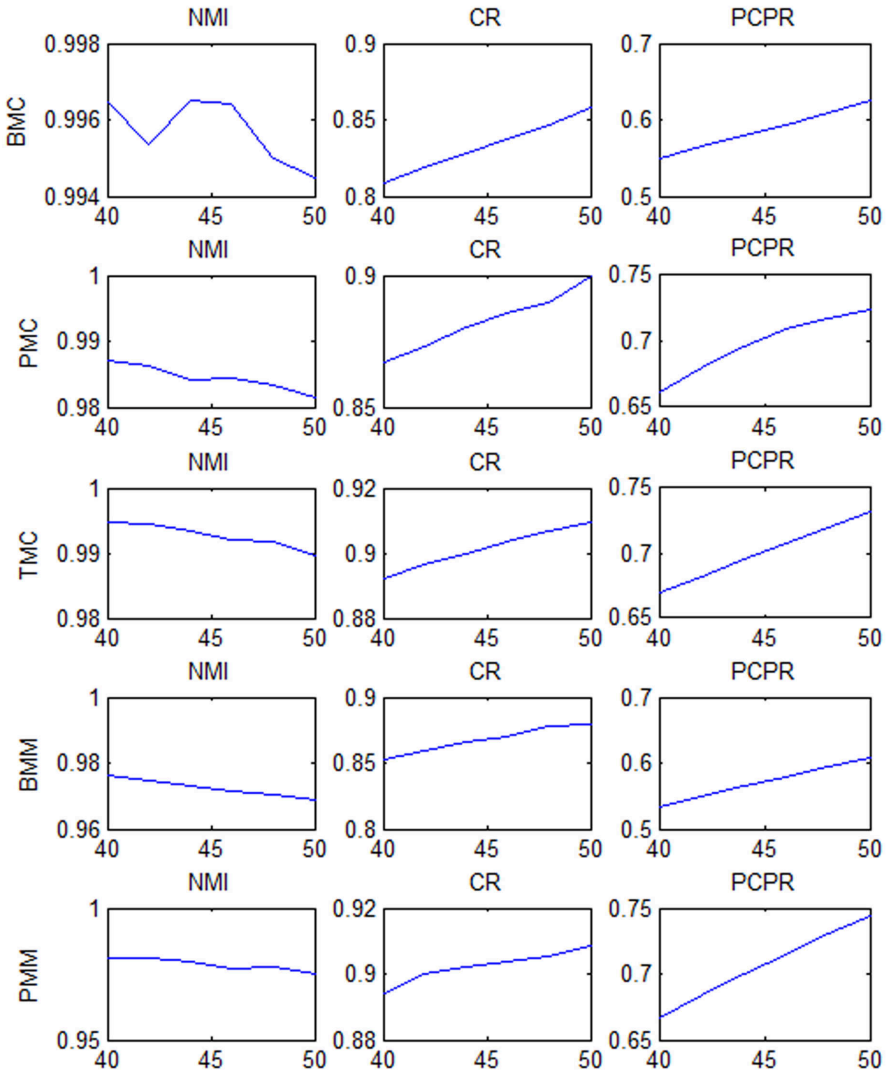

Figure 5.23: Cost functions for the range of 40 to 50 for $\sigma = 16\%$ for the tested images using NMI [27], CR [34], and the proposed PCPR objective function.

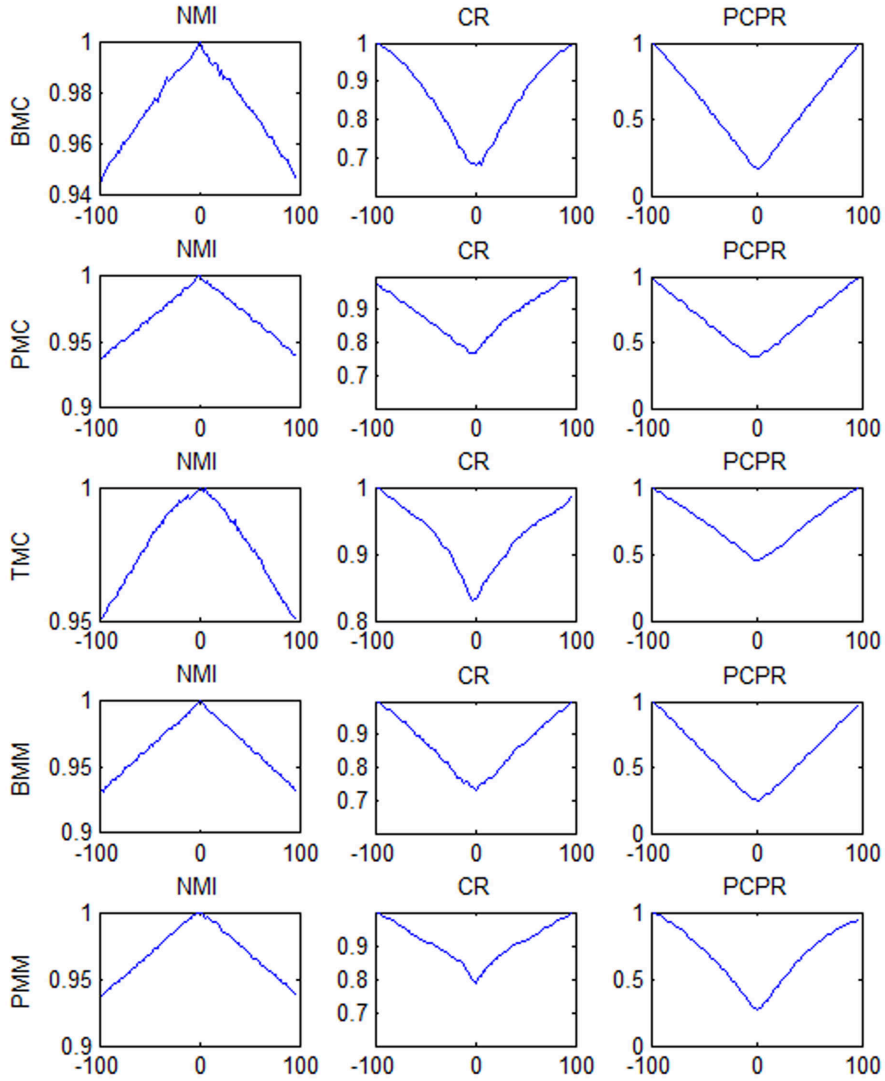

Figure 5.24: Cost functions for the range of -100 to 100 for $\sigma = 22\%$ for the tested images using NMI [27], CR [34], and the proposed PCPR objective function. NMI performs very poorly under this high noise scenario, exhibiting noticeable local optima and increasing dissimilarity as the images come into alignment. The cost functions produced by CR now exhibit local optima near the true alignment for the BMC case. As with the previous noise levels, the cost function produced by PCPR remain largely smooth and converge to the true alignment of the images.

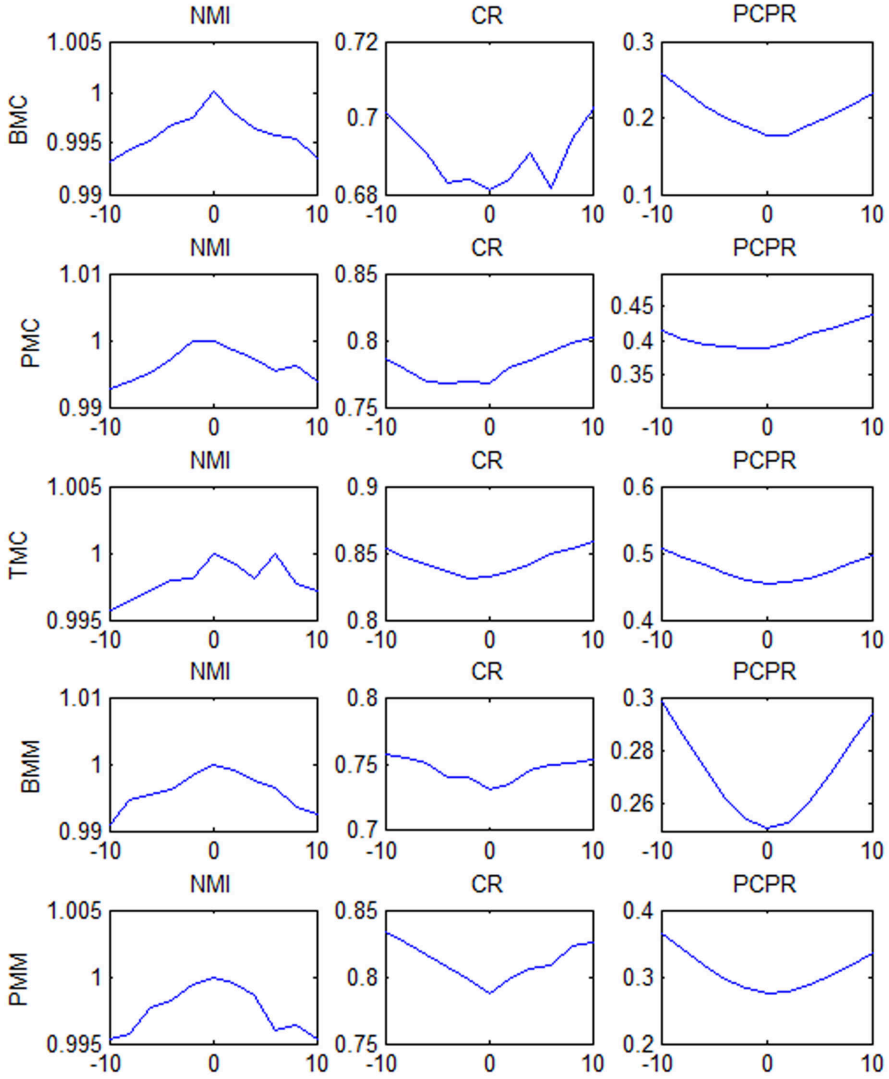

Figure 5.25: Cost functions for the range of -10 to 10 for $\sigma = 22\%$ for the tested images using NMI [27], CR [34], and the proposed PCPR objective function. NMI performs very poorly under this high noise scenario, exhibiting noticeable local optima and increasing dissimilarity as the images come into alignment. The cost functions produced by CR now exhibit local optima near the true alignment for the BMC case, in addition to the local optima present in the BMM case. As with the previous noise levels, the cost function produced by PCPR remain largely smooth and converge to the true alignment of the images.

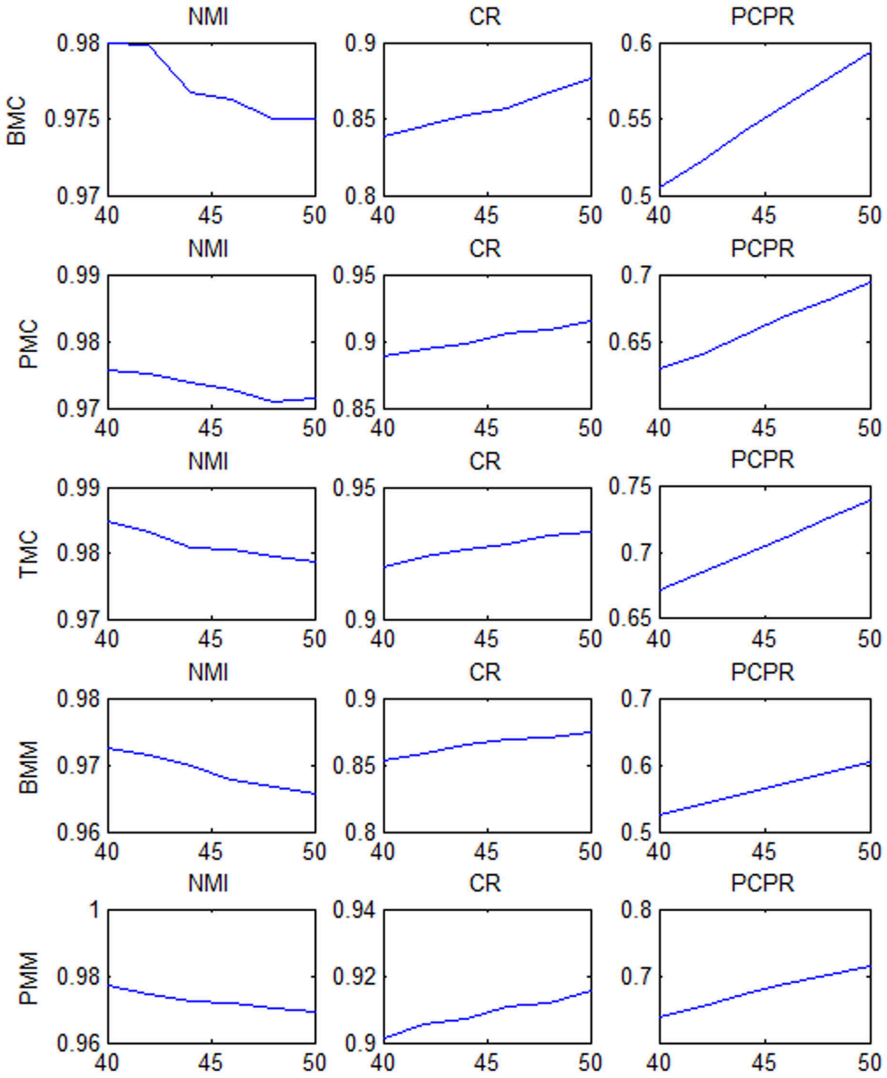

Figure 5.26: Cost functions for the range of 40 to 50 for $\sigma = 22\%$ for the tested images using NMI [27], CR [34], and the proposed PCPR objective function.

functions that are largely smooth with global optima that corresponds with the ground-truth alignment between the tested images under contrast non-uniformities and high levels of noise. Given the proposed objective function, the next step is to study the performance of the proposed objective function for multimodal image registration under different scenarios.

Chapter 6

Registration Results

Given the PCPR objective function proposed in Chapter 5, the next step is to illustrate the registration performance of the proposed PCPR objective function for multimodal image registration under different scenarios. This chapter provides a performance analysis of the proposed objective function under different scenarios. First, the registration accuracy using the PCPR objective function was evaluated under different 2D registration scenarios to assess performance more comprehensively under controlled situations. Second, the registration accuracy using the PCPR objective function was evaluated under different volumetric registration scenarios to better assess performance under complex scenarios.

6.1 2D Registration

To facilitate for a more comprehensive performance analysis, different 2D registration scenarios were performed using the test image pairs used in Section 5.6. First, registration accuracy using the PCPR objective function is evaluated under ideal conditions. Second, registration accuracy using the PCPR objective function is evaluated under scenarios characterized by contrast non-uniformities to evaluate its robustness to contrast non-uniformities. Finally, registration accuracy is evaluated under different noise contaminated scenarios to evaluate its robustness to noise.

6.1.1 Registration under Ideal Conditions

To study the registration accuracy under ideal conditions (e.g., no noise or contrast non-uniformities), the test image pairs used in Section 5.6 were distorted and subsequently registered using the proposed PCPR objective function described in Eq. (5.19). For comparison purposes, the NMI [27] and CR [34] objective functions were also tested. A simple gradient descent optimization scheme is utilized to determine the transformation that aligns the images together based on these objective functions to allow for a fair comparison between the tested methods. Each of the test image pairs were subjected to the following tests:

1. horizontal translations of -80 and 80 pixels, and
2. rotations of -40 and 40 degrees.

This creates a total of 20 test cases. Registration accuracy for all objective functions are evaluated quantitatively based on the translation registration error and the rotation registration error for the translation and rotation distortion cases respectively. The translation registration error is defined as the difference between the alignment translation determined using the objective function and the ground-truth alignment on a pixel basis. The rotation registration error is defined as the difference between the alignment rotation determined using the objective function and the ground-truth alignment on a degree basis.

Histograms of the translation and rotation registration errors for all six test image sets are shown in Fig. 6.1 and Fig. 6.2 respectively. In Fig. 6.1, all three objective functions provide low translation registration errors with relatively low error variance. However, the PCPR objective function provides the lowest mean translation registration error when compared to the NMI and CR objective functions. In Fig. 6.2, both the NMI and PCPR objective functions provide zero rotation registration errors in all test cases. Unfortunately, the CR objective function showed high rotation registration errors as well as a high error variance. This indicates that the optimization method becomes stuck in local optima in the CR objective function and is due to the fact that the functional mapping between the images used by the CR objective function can vary greatly as we converge to the global optima.

The registration results for the BMM case distorted by a horizontal translation of 80 and the PMM case distorted by a clockwise rotation of 40^o are shown in Fig. 6.3 and Fig. 6.4 respectively. By visual inspection, all three objective functions allowed for good

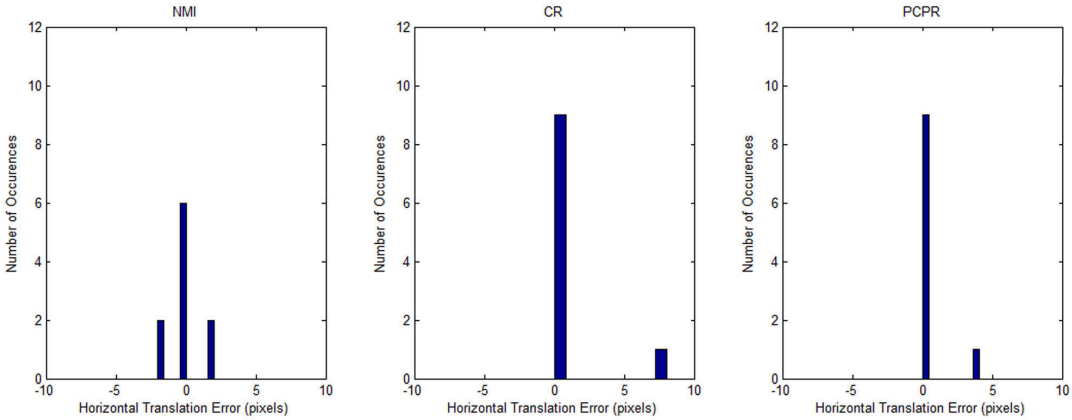

Figure 6.1: Histogram of registration error for horizontal translations of -80 and 80 for all test sets.

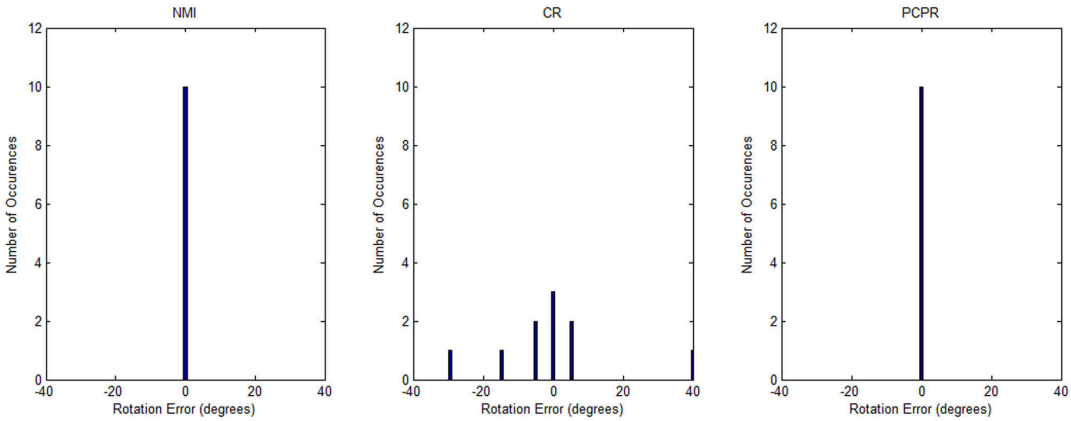

Figure 6.2: Histogram of registration error for rotations of 40^o and 40^o for all test sets.

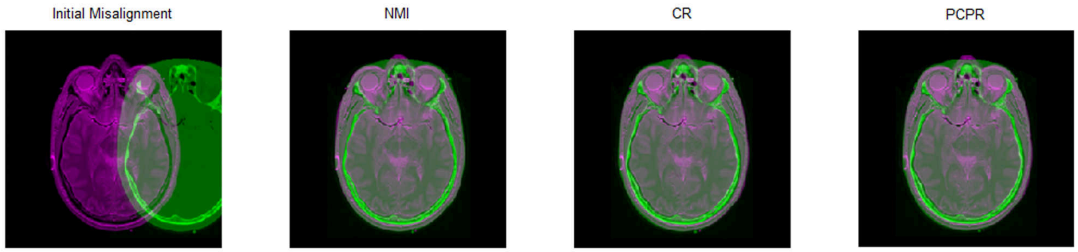

Figure 6.3: Registration results for the BMC case distorted by a horizontal translation of 80.

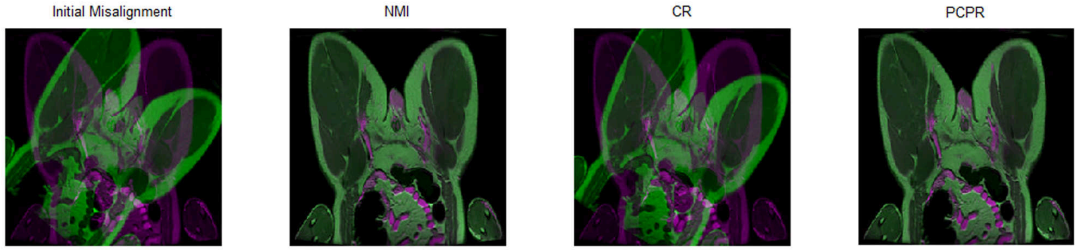

Figure 6.4: Registration results for the PMM case distorted by a clockwise rotation of 40^o.

registration accuracy for the BMM case, with the exception of minor misalignment in the results produced by the NMI objective function. In the PMM case, both the NMI and PCPR objective functions provided good registration accuracy while the results produced by the CR objective function remains noticeably misaligned. These results show that in general both the NMI and PCPR objective functions perform well under ideal scenarios, while the CR objective function performs poorly in the rotation distortion scenarios.

6.1.2 Registration under Non-uniformity

To study the registration accuracy under contrast non-uniformities, the test image pairs with contrast non-uniformities used in Section 5.5 were distorted and subsequently registered using the NMI, CR, and PCPR objective functions. Each of the test image pairs were subjected to the same distortions as that used in Section 6.1.1 for a total of 20 test cases. Registration accuracy using all objective functions are evaluated quantitatively based on the translation registration error and the rotation registration error for the translation and rotation distortion cases respectively.

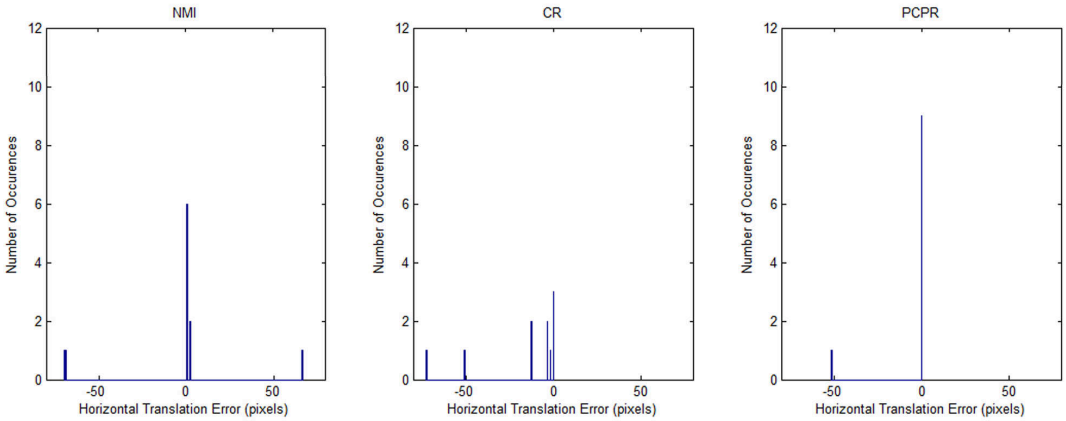

Figure 6.5: Histogram of registration error under contrast uniformities for horizontal translations of -80 and 80 for all test sets.

Histograms of the translation and rotation registration errors for all six test image sets are shown in Fig. 6.5 and Fig. 6.6 respectively. In Fig. 6.5, the proposed method provides the lowest mean translation registration error and error variance when compared to the NMI and CR objective functions. In Fig. 6.6, the NMI and PCPR objective functions provide zero translation registration error for all of the test cases. Unfortunately, the CR objective function showed the highest translation and rotation registration errors as well as a high error variance. This indicates that the optimization scheme becomes stuck in local optima in the CR objective function and is due to the fact that the local contrast non-uniformities result in poor functional mappings between the images used by the CR objective function.

The registration results for the BMM case distorted by a horizontal translation of -80 and the BMC case distorted by a counterclockwise rotation of 40^o under contrast uniformities are shown in Fig. 6.7 and Fig. 6.8 respectively. By visual inspection, the NMI and CR objective functions both performed poorly for the BMM case, with the PCPR objective function producing good alignment between the images. In the BMC case, both the NMI and proposed objective functions provided good registration accuracy while the results produced by the CR objective function remains noticeably misaligned. These results show that the PCPR objective function provides good registration accuracy in scenarios characterized by contrast non-uniformities.

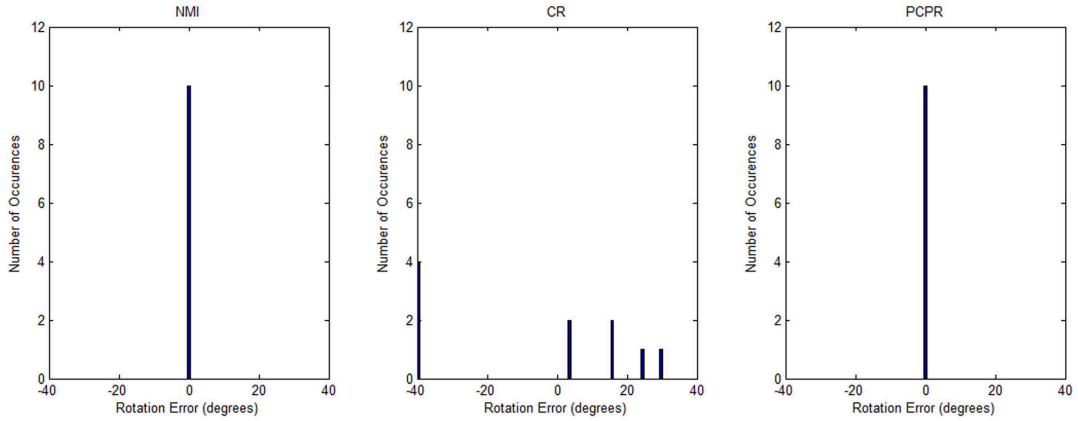

Figure 6.6: Histogram of registration error under contrast uniformities for rotations of 40^o and 40^o for all test sets.

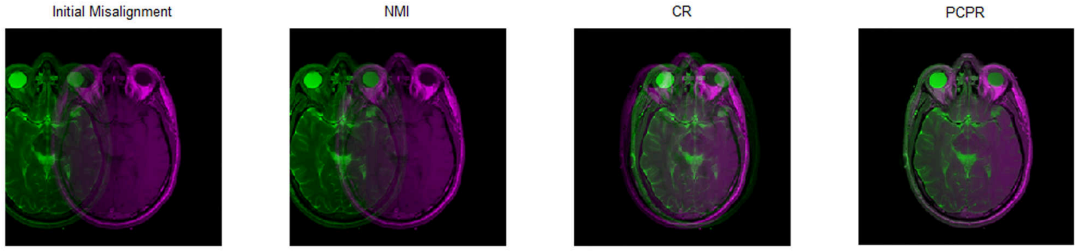

Figure 6.7: Registration results for the BMM case under contrast uniformities distorted by a horizontal translation of -80.

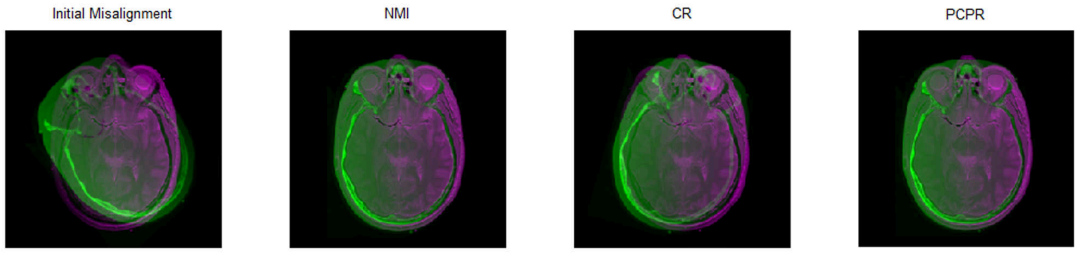

Figure 6.8: Registration results for the BMC case under contrast uniformities distorted by a counterclockwise rotation of 40^o.

6.1.3 Registration under Noise

To study the registration accuracy under different levels of noise, the test image pairs used in Section 5.6 were contaminated with additive Gaussian noise with standard deviations of $\sigma = \{8\%, 16\%, 22\%\}$, distorted, and subsequently registered using the NMI, CR, and PCPR objective functions. Each of the test image pairs were subjected to same translation distortions as that used in Section 6.1.1 for a total of 30 test cases. Registration accuracy for all methods are evaluated quantitatively based on the translation registration error and the rotation registration error for the translation and rotation distortion cases respectively.

Histograms of the translation registration errors for all six test image sets for the different noise scenarios are shown in Fig. 6.9, Fig. 6.10, and Fig. 6.11. The PCPR objective functions achieved the lowest mean translation registration errors and error variances under all of the noise scenarios. The CR objective function achieved noticeably higher mean translation registration errors and error variances when compared with the PCPR objective function, but is still able to achieve translation registration error within 2 pixels in at least 60% of the cases. The NMI objective function had the highest mean translation errors when compared with the CR and NMI objective functions, with zero cases where the translation registration error is within 2 pixels at noise levels of $\sigma = \{16\%\}$ and $\sigma = \{22\%\}$. This is due to the fact that the joint and marginal intensity distributions computed in the presence of noise is not representative of the statistical characteristics of the underlying image content. Since the computation of NMI relies on reliable joint and marginal intensity distributions, the NMI objective function is highly sensitive to the presence of noise.

The registration results for the PMC case distorted by a horizontal translation of -80 under additive Gaussian noise with $\sigma = 8\%$ is shown in Fig. 6.12. By visual inspection, the registration results produced by the NMI objective function is noticeably misaligned, with the CR objective function producing results with a minor misalignment and the proposed method producing the correct alignment between the images. The registration results for the BMM case distorted by a horizontal translation of 80 under additive Gaussian noise with $\sigma = 16\%$ is shown in Fig. 6.13. Similarly, the registration results produced by the NMI objective function is noticeably misaligned. Fortunately, the CR and PCPR objective functions produced the correct alignment between the images. Finally, the registration results for the TMC case distorted by a horizontal translation of -80 under additive Gaussian noise with $\sigma = 22\%$ is shown in Fig. 6.14. The registration results produced by the NMI objective function is noticeably misaligned, with the CR objective function producing results with a minor misalignment and the proposed method producing the correct alignment between the images. These results show that the PCPR objective function provides good

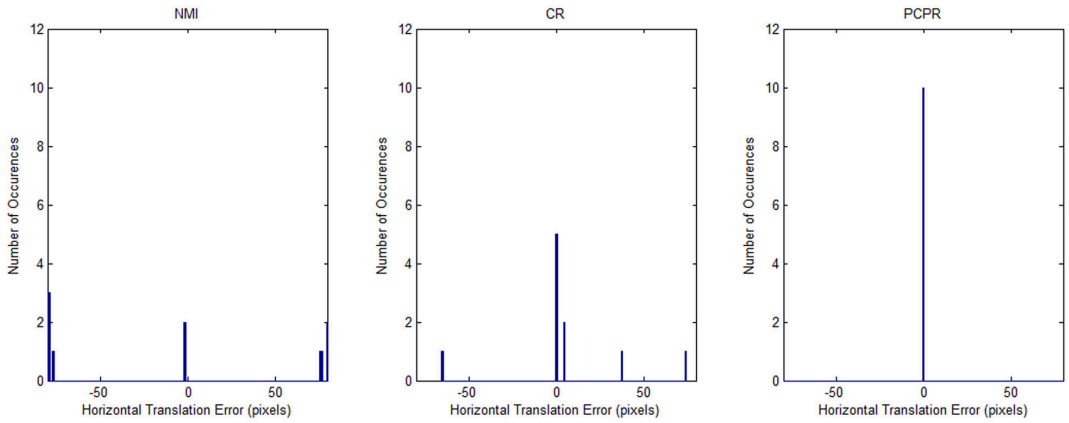

Figure 6.9: Histogram of registration error for $\sigma = 8\%$ for horizontal translations of -80 and 80 for all test sets.

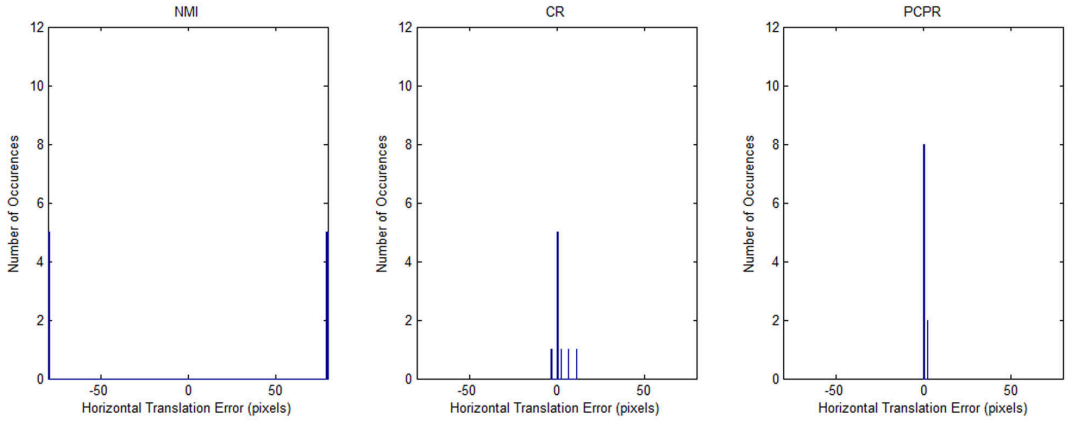

Figure 6.10: Histogram of registration error for $\sigma = 16\%$ for horizontal translations of -80 and 80 for all test sets.

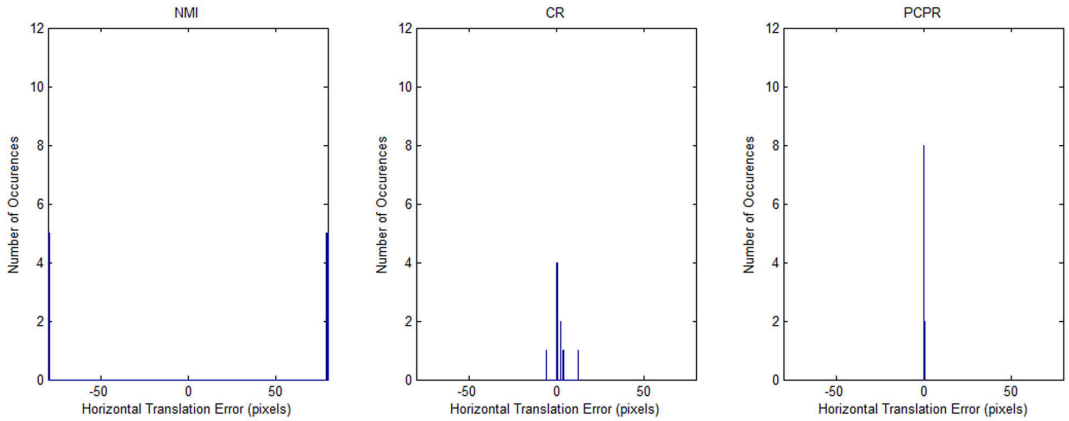

Figure 6.11: Histogram of registration error for $\sigma = 22\%$ for horizontal translations of -80 and 80 for all test sets.

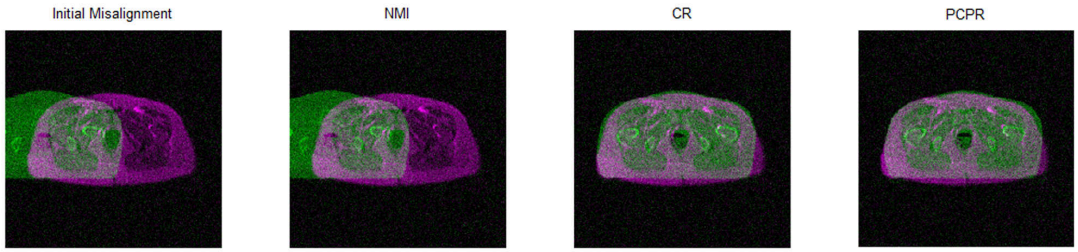

Figure 6.12: Registration results for the PMC case for $\sigma = 8\%$ distorted by a horizontal translation of -80.

registration accuracy in scenarios characterized by high noise levels.

6.2 Volumetric Registration

To study the performance of the proposed PCPR objective function under more complex scenarios, volumetric registration using 9 real patient MR-CT brain volume data sets from the Whole Brain Atlas [90] (WBA) was performed under different scenarios. The MR volumes are $256 \times 256 \times 23$ voxels, with a slice thickness of 5mm. The CT volumes are downsampled to the same voxel dimensions. A summary of each test data set is given below.

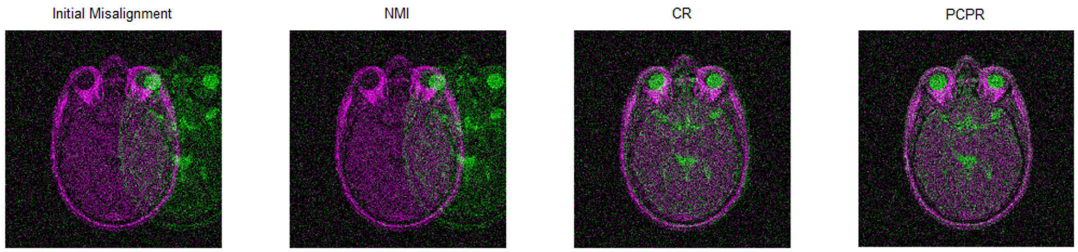

Figure 6.13: Registration results for the BMM case for $\sigma = 16\%$ distorted by a horizontal translation of 80.

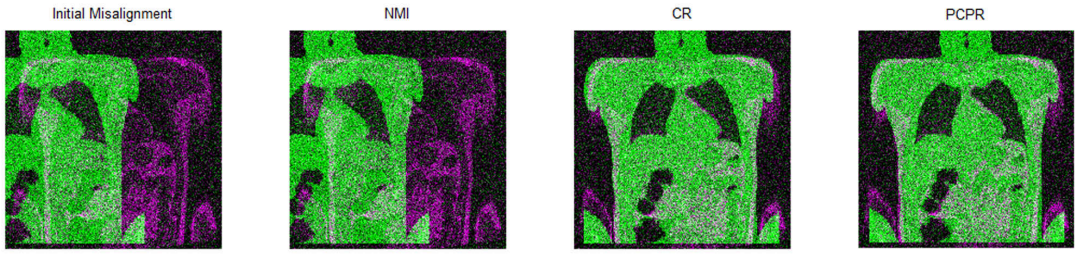

Figure 6.14: Registration results for the TMC case for $\sigma = 22\%$ distorted by a horizontal translation of -80.

1. **Test 1**: MR/PD-CT, 63 year-old male.
2. **Test 2**: MR/T2-CT, 63 year-old female.
3. **Test 3**: MR/T2-CT, 45 year-old female.
4. **Test 4**: MR/T2-CT, 23 year-old female.
5. **Test 5**: MR/PD-CT, 42 year-old female.
6. **Test 6**: MR/PD-CT, 75 year-old male.
7. **Test 7**: MR/T2-CT, 22 year-old male.
8. **Test 8**: MR/T2-CT, 55 year-old male.
9. **Test 9**: MR/T2-CT, 71 year-old female.

For testing the objective functions, the following volumetric registration algorithm was used. Given the optimization problem for multimodal image registration described in Eq. (5.1), a sequential quadratic programming approach [17] is then employed to solve the problem, where the estimated transformation $\hat{T}$ at iteration k can be defined as

$$\hat{T}_k = \hat{T}_{k-1} + \gamma_{k-1} d_{k-1}, \tag{6.1}$$

where γ is a non-negative step size and d is the step direction calculated by solving a quadratic subprogram involving Ψ [17]. The pseudo-code for registering two volumes f and g is presented in Algorithm 2.

Algorithm 2 Volumetric Registration

Require: Given $T_0 = I$, $k = 0$
1: **repeat**
2: Compute similarity $\Psi(f(T_k(\underline{x})), g(\underline{x}))$.
3: $k = k + 1$.
4: Estimate $\hat{T}_k$ (6.1).
5: **until** $\triangle\Psi < \epsilon_{convergence}$
6: **return** $\hat{T}_k$.

For evaluation purposes, the NMI and PCPR objective functions were tested.

6.2.1 Registration under Noise-free Conditions

The first set of tests studies the registration accuracy using NMI and PCPR under geometric distortion scenarios. Each test data set was distorted using 30 randomly generated affine transformations, based on the random perturbation of translation coefficients up to ±30mm and all other coefficients up to ±0.1, resulting in a total of 270 test cases. Since the test image sets used were previously aligned, the gold-standard transformations are known for all 270 test cases. Registration accuracy for all methods is evaluated quantitatively based on the fiducial registration error (FRE), which in our case can be defined by the root-mean-square error of 60 fiducial points within the region of interest. The fiducial points were chosen randomly within the regions of interest in the test images as not to bias the tests towards any of the tested similarity measures.

The registration results for all nine test data sets, totalling 270 test cases, are summarized in Table 6.1. PCPR was capable of achieving noticeably lower FRE when compared to NMI for all test cases. One contributing factor to this difference in registration error when compared to NMI is that the intensity relationships between the tested MR and CT volume data sets are highly complex and nonlinear, making NMI highly non-convex and difficult to optimize. On the other hand, the structural relationships between the volume data sets is significantly more straightforward, making PCPR more straightforward to optimize. Sample registration results using NMI and PCPR for Test 1 shown in Fig. 6.15. Visually, both NMI and PCPR are capable of providing accurate registration results. These experimental results demonstrate the effectiveness of the PCPR objective function for performing non-rigid multimodal image registration on CT and MR images.

6.2.2 Registration under Noise

An important consideration in the design of an objective function for image registration is that the objective function should be robust to noise artifacts that can affect registration accuracy. To study the effect of noise on the similarity measures, the set of tests conducted in Section 6.2.1 for Test 1 were performed with the MR volumes contaminated by 5%, 10%, 15%, and 20% simulated Rician noise, resulting in a total of 120 tests.

The registration results for the Test 1 under the various noise levels, totalling 120 test cases, are summarized in Table 6.2. The FRE achieved using PCPR remained largely consistent at all noise levels. On the other hand, the FRE achieved using NMI rose significantly as noise levels increased. Sample registration results using NMI and PCPR for

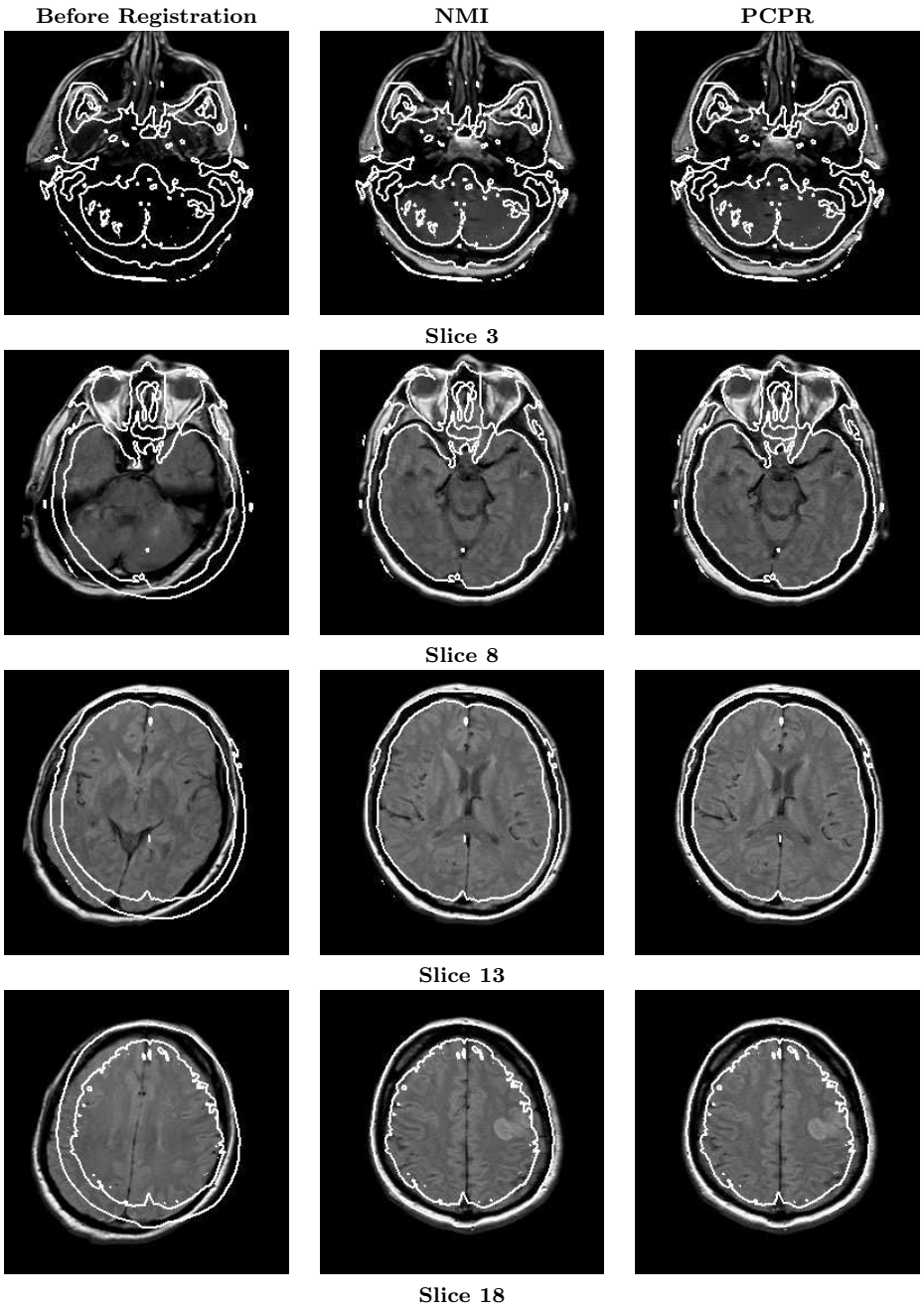

Figure 6.15: Sample registration result from Test 1 using NMI and PCPR. For visualization, slices from the volumes are shown. Contours extracted from the CT volume are overlayed on the MR volume to visualize the quality of registration. Visually, the MR volumes warped based on PCPR and NMI appear well aligned with the CT volume.

Table 6.1: Fiducial registration errors (FRE) of NMI and PCPR for the 9 data sets. A total of 30 random distortions is tested for each data set, for a total of 270 test cases.

Test Set	**FRE (mm)**	
	NMI	**PCPR**
Test 1	3.74±0.56	**2.21**±0.29
Test 2	3.91±0.50	**2.09**±0.34
Test 3	2.36±0.63	**2.31**±0.43
Test 4	3.64±0.49	**2.08**±0.27
Test 5	3.81±0.30	**2.31**±0.22
Test 6	3.71±0.47	**2.40**±0.28
Test 7	2.90±0.53	**2.57**±0.20
Test 8	3.35±0.55	**2.39**±0.23
Test 9	3.97±0.61	**2.12**±0.35

20% simulated Rician noise are shown in Fig. 6.16. Visually, PCPR is capable of providing accurate registration results, while the results produced by NMI appears misaligned. These experimental results demonstrate the robustness of PCPR to the presence of noise artifacts.

Table 6.2: Fiducial registration errors (FRE) of NMI and PCPR for Test 1 under different Rician noise levels for MR volume. A total of 30 random distortions is tested for each noise level, for a total of 120 test cases.

	FRE (mm)			
Rican noise	**5%**	**10%**	**15%**	**20%**
NMI	4.02±0.76	6.73±1.03	9.71±1.64	10.86±2.58
PCPR	2.24±0.35	2.28±0.37	2.34±0.59	2.53±0.62

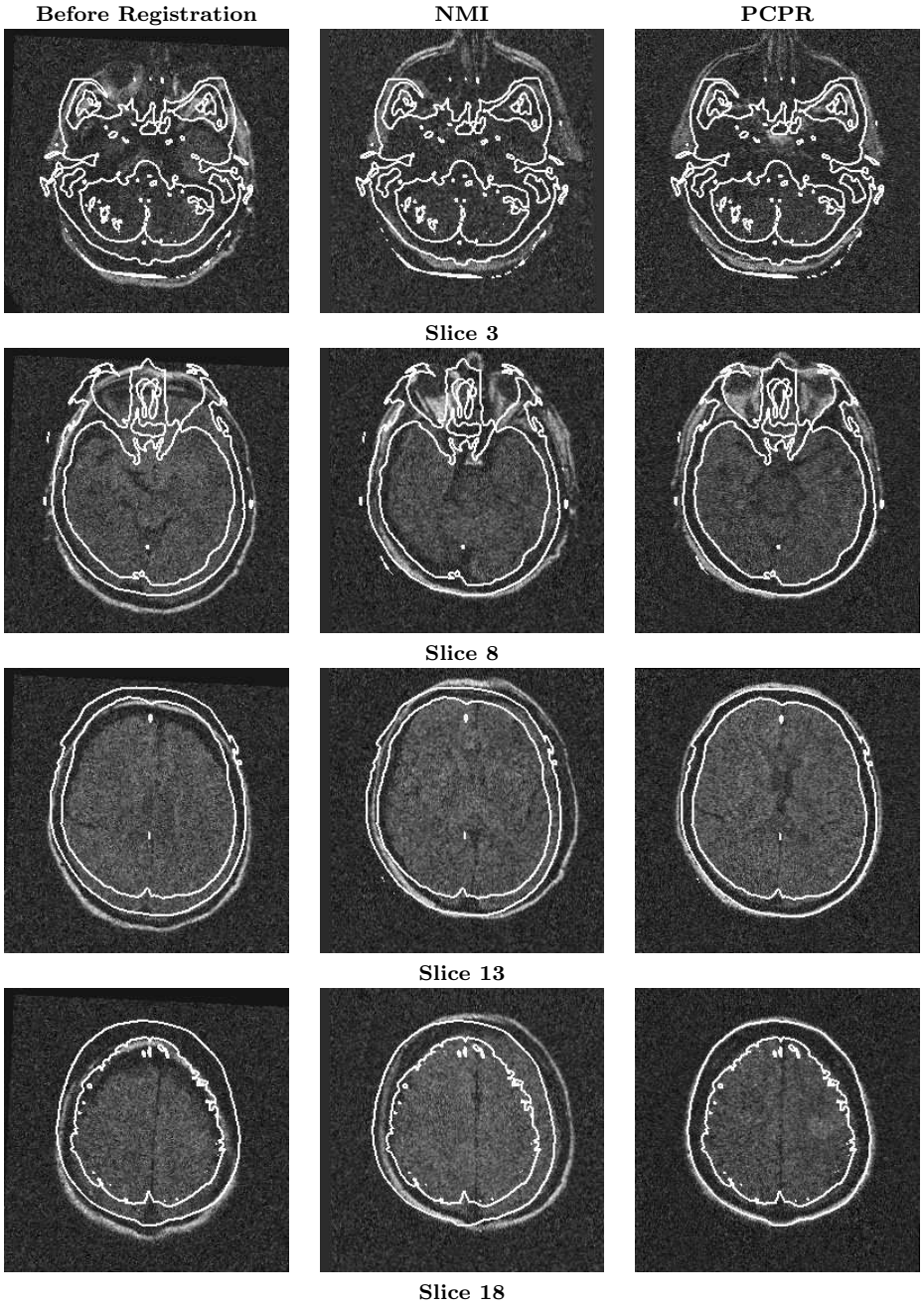

Figure 6.16: Sample registration result from Test 1 using NMI and PCPR under 20% Rician noise. For visualization, corresponding slices from the volumes are shown. Contours extracted from the CT volume are overlayed on the MR volume to visualize the quality of registration. Visually, the MR volume warped based on PCPR appears well aligned with the CT volume despite noise, while the results produced by NMI appears misaligned.

Chapter 7

Conclusions

The thesis dissertation presents a robust objective function for registering images acquired using different imaging modalities. The results of this study demonstrate the improved registration accuracies under different imaging conditions can be gained using the proposed PCPR objective function when compared to existing objective functions. This thesis provides a foundation for future research in automatic multimodal image registration by making several contributions.

7.1 Summary of contributions

The contributions of the dissertation include:

- introduces a novel, robust image representation based on complex phase relationships,
- introduces a novel multi-scale representation based on probability theory that provides improved structural preservation and localization at coarse scales, as well as improved noise robustness,
- introduces the concept of probabilistic complex phase representation as the basis for multimodal image registration,
- provides experimental results demonstrating the effectiveness of the proposed probabilistic complex phase representation at providing improved structural localization and reduced noise sensitivity when compared to existing linear and nonlinear scale space representations,

- proposes an objective function for performing multimodal image registration based on a Geman-McClure error distribution model between probabilistic complex phase representations, and
- provides experimental results demonstrating the effectiveness of the proposed objective function in achieving improved registration accuracies under non-uniformity and noise conditions when compared to existing objective functions.

7.2 Parallel Research and Impact

While the concept of probabilistic image representations has been investigated in this dissertation for the purpose of multimodal image registration, the impact of these concepts span several different areas of image processing and computer vision research. In particular, the author has applied this concept to the following areas:

Illumination and reflectance separation [91] : A probabilistic sampling approach was introduced for learning and estimating the underlying illumination from an image, which in turn was used to estimate the reflectance. This approach takes advantage of both local and global contrast information to provide better separation of reflectance and illumination by reducing the effects of strong shadows and other sharp illumination changes on the estimation process.

Image synthesis [92] : Assuming a Markov random field model, a probabilistic sampling approach was introduced for learning the conditional posterior distribution of SAR sea ice data on a regional basis. Based on the learned model, synthetic SAR sea-ice data and the associated ground-truth segmentations was generated for performing systematic and reliable objective evaluation of SAR sea-ice image segmentation methods.

Edge detection [93] : A probabilistic image model was introduced for decomposing images into multi-scale representations that are robust to noise and provide strong structural preservation. Based on this model, an edge detection method was introduced that provided strong structural extraction even under high noise scenarios.

Blind decorrelation [94] : A probabilistic sampling approach was introduced for learning the non-stationary point spread function directly from the SAR data using a

Fisher-Tippett scatter model. Based on learned point spread function, blind decorrelation can then be performed on the SAR data to facilitate for improved speckle reduction.

Keypoint extraction [95] : A probabilistic multi-scale approach was introduced for the purpose of robust keypoint extraction in high-noise environments. After decomposing the image using a probabilistic sampling approach, a gradient second-order moment analysis approach was employed to identify initial keypoint candidates. The final keypoints and their characteristic scales are determined based on the local Hessian trace extrema over all scales. By taking advantage of the structural localization and noise robustness of the multi-scale decomposition, the keypoint extraction approach was shown to be highly effective at finding unique keypoints under heavy noise.

Image restoration [96, 97] : A probabilistic sampling approach was introduced for learning a probabilistic model of the image. Based on the learned probabilistic model, a Bayesian estimation approach was employed to estimate the original image.

Similarly, the concept of complex phase representations investigated in this dissertation has a broad impact on several different areas of image processing and computer vision research. In particular, the author has applied this concept to the following areas:

Super-resolution [98] : The super-resolution problem was formulated as a constrained optimization problem using third-order Markov prior model, and the novel approach for adapting the priors based on the phase variations of the low-resolution images was introduced for reconstructing the high-resolution images. This approach was shown to be effective at preserving visual information in the reconstructed high-resolution images.

Noise reduction [99] : A phase-adaptive bilateral filtering approach was introduced for suppressing noise in images. By adapting the constraints of the bilateral kernel based on the underlying phase information, this approach was shown to be effective at suppressing image noise while preserving structural information.

Image fusion [100] : A phase-adaptive approach was introduced for fusing different images together into a single image. By adapting the contribution of information from different images based on their phase characteristics, this approach can improve the visualization of important characteristics from different images in the fused image.

Image segmentation [101] : An active contour segmentation approach was introduced where phase moment information extracted from the image is used as the external energy functional. By employing phase moment information, this approach allowed for robust segmentation under illumination variation scenarios.

7.3 Future Research

The research presented in this dissertation provides a foundation for future research in automatic multimodal image registration. Three potential research topics that can build upon the research presented in this dissertation are presented in the following sections.

7.3.1 Stochastic Optimization

For the purpose of this research, we have primarily focused on the design of the objective function used to determine the alignment between two images. As such, for the optimization process used to determine the alignment between two images, we have mainly employed deterministic optimization approaches. While testing has shown that the cost function produced by the proposed objective function is reasonably smooth and monotonic compared to other objective functions for 2D cases, local optima may exist for higher-dimensional data as well as for more complex geometric distortions. Since deterministic optimization approaches have a tendency to be trapped in such local optima, as future work, it is worth investigating the development of stochastic optimization approaches that are more robust to such local optima.

7.3.2 Joint Multimodal Registration

The proposed objective function is designed for registering multimodal images in a pair-wise manner. A problem with doing pair-wise registrations to register multiple images is that the resulting alignment is typically sub-optimal for the entire set of images since internal consistencies are not maintained between all the images. Therefore, a better approach to the problem of registering multiple images from different modalities may be to extend the proposed objective function such that it can be used to register all images in a simultaneous fashion to improve internal consistency between the images, which will be investigated in future work.

7.3.3 Generalized Probabilistic Framework for Signal Representation

The probabilistic image representation presented in this dissertation marks the beginning in the development of a generalized probabilistic framework for discrete signal representation. By representing a signal not as a set of fixed values (e.g., intensity) but as a set of discrete probability distribution functions (e.g., one probabilistic distribution function per time instance or pixel site) estimated using probabilistic sampling approaches, statistical techniques can be employed to facilitate for robust signal processing and computer vision tasks such as signal restoration, signal enhancement, signal reconstruction, source separation, feature extraction, motion tracking and analysis, image synthesis, segmentation, registration, and classification.

References

[1] M. Pickering, J. Arnold, and M. Frater. An adaptive search algorithm for block matching motion estimation. *IEEE Transactions on Circuits and Systems for Video Technology*, 7(6):906–912, 1997. 1

[2] C. Cheung and L. Po. Adjustable partial distortion search algorithm for fast block motion estimation. *IEEE Transactions on Circuits and Systems for Video Technology*, 13(1):100–110, 2003. 1

[3] C. Hsu, Y. Wu, and A. Chen. Content-based image retrieval by feature point matching. In *Proceedings of SPIE, Storage and Retrieval for Media Databases*, volume 4315, pages 39–49, 2001. 1

[4] C. Hsu and M. Shih. Content-based image retrieval by interest-point matching and geometric hashing. In *Proceedings of SPIE, Electronic Imaging and Multimedia Technology III*, volume 4925, pages 80–90, September 2002. 1

[5] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multi-modality image registration based on information theory. In *Proceedings of international conference on information processing in medical imaging*, pages 263–274, 1995. 1, 5, 7

[6] P. Viola and W. Wells. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):137–154, 1997. 5, 7

[7] W. Li and H. Leung. A maximum likelihood approach for image registration using control point and intensity. *IEEE Transactions on Image Processing*, 13(8):1115–1127, 2004. 7

[8] H. Chen, P. Varshney, and M. Arora. Mutual information based image registration for remote sensing data. *International Journal of Remote Sensing*, 24(18):3701–3706, 2003. 1, 5, 7

[9] L. Hong and A. Jain. Integrating faces and fingerprints for personal identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12):1295–1307, 1998. 1

[10] D. Mattes, David R. Haynor, H. Vesselle, Thomas K. Lewellen, and W. Eubank. PET-CT image registration in the chest using free-form deformations. *IEEE Transactions on Medical Imaging*, 22(1):120–128, 2003. 1

[11] L. Chen, T. Teo, J. Rau, J. Liu, and W. Hsu. Building reconstruction from lidar data and aerial imagery. In *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, pages 2846–2849, 2005. 1

[12] G. Vosselman. Fusion of laser scanning data, maps and aerial photographs for building reconstruction. In *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, volume 1, pages 85–88, 2002. 1

[13] A. Simmons, P. Tofts, G. Barker, and S. Arridge. Sources of intensity nonuniformity in spin echo images at 1.5t. *Magnetic Resonance in Medicine*, 32(1):121–128, 1994. 3

[14] M. Oghabian, S. Mehdipour, and N. Alam. The impact of rf inhomogeneity on mr image non-uniformity. In *Proceedings of Image and Vision Computing New Zealand*, 2003. 3

[15] R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49(6):409–436, 1952. 3

[16] J. Nelder and R. Mead. A simplex method for function minimization. *Computer Journal*, 7:308–313, 1965. 3

[17] P. Boggs and J. Tolle. Sequential quadratic programming. *Acta Numerica*, pages 1–51, 1995. 3, 107

[18] S. Ranade and A. Rosenfeld. Point pattern matching by relaxation. *Pattern Recognition*, 12:269–275, 1980. 5, 6

[19] C. Wang, H. Sun, S. Yada, and A. Rosenfeld. Some experiments in relaxation image matching using corner features. *Pattern Recognition*, 16:167–182, 1983.

[20] J. Ton and A. Jain. Registering landsat images by point matching. *IEEE Transactions on Geoscience and Remote Sensing*, 27(5):642–651, 1989. 6

[21] Y. Chen and G. Medioni. Object modelling by registration of multiple range images. *Image Vision Computing*, 10(3):145–155, 1992. 6

[22] S. Rusinkiewicz N. Gelfand, L. Ikemoto and M. Levoy. Geometrically stable sampling for the icp algorithm. In *Proceedings of the International Conference on 3D Digital Imaging and Modeling*, 2003.

[23] T. Jost and H. Hugli. A multi-resolution icp with heuristic closest point search for fast and robust 3d registration of range images. In *Proceedings of the International Conference on 3D Digital Imaging and Modeling*, 2003. 5, 6

[24] E. Castro and C. Morandi. Registration of translated and rotated images using finite fourier transforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5):700–703, 1987. 5, 6

[25] B. Reddy and B. Chatterji. An fft-based technique for translation, rotation and scale invariant image registration. *IEEE Transactions on Image Processing*, 5(8):1266–1271, 1996. 6

[26] I. Zavorin and J. Le Moigne. Use of multiresolution wavelet feature pyramids for automatic registration of multisensor imagery. *IEEE Transactions on Image Processing*, 14(6):770–782, 2005. 5, 6

[27] C. Studholme, D. Hill, and D. Hawkes. An overlap invariant entropy measure of 3d medical image alignment. *Pattern Recognition*, 32(1):71–86, 1999. 5, 7, 22, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 84, 85, 87, 88, 89, 90, 91, 92, 93, 94, 95, 98

[28] A. Fitch, A. Kadyrov, W. Christmas, and J. Kittler. Fast robust correlation. *IEEE Transactions on Image Processing*, 14(8):1063–1073, 2005. 7

[29] J. Kybic and M. Unser. Fast parametric elastic image registration. *IEEE Transactions on Image Processing*, 12(11):1427–1442, 2003.

[30] Z. Li, Z. Bao, H. Li, and G. Liao. Image autocoregistration and insar interferogram estimation using joint subspace projection. *IEEE Transactions on Geoscience and Remote Sensing*, 44(2):288–297, 2006.

[31] J. Orchard. Efficient global weighted least-squares translation registration in the frequency domain. In *Proceedings of the International Conference on Image Analysis and Recognition*, pages 116–124, 2005. 7

[32] A. Refice, F. Bovenga, and R. Nutricato. Mst-based stepwise connection strategies for multipass radar data, with application to coregistration and equalization. *IEEE Transactions on Geoscience and Remote Sensing*, 44(8):2029–2040, 2006.

[33] A. Roche, G. Malandain, X. Pennec, and N. Ayache. The correlation ratio as a new similarity measure for multimodal image registration. In *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention", volume = "1496", year = "1998", pages = "1115–1124",*. 7

[34] A. Roche, X. Pennec, M. Rudolph, D. P. Auer, G. Mal, S. Ourselin, L. M. Auer, and N. Ayache. Generalized correlation ratio for rigid registration of 3d ultrasound with mr images. In *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention*, volume 1935. 8, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 84, 85, 87, 88, 89, 90, 91, 92, 93, 94, 95, 98

[35] A. Guimond, A. Roche, N. Ayache, and J. Meunier. Three-dimensional multimodal brain warping using the demons algorithm and adaptive intensity corrections. *IEEE Transactions on Medical Imaging*, 20(1):58–69, 2001. 7

[36] M. Omanovic and J. Orchard. Efficient multimodal registration using least-squares. In *Proceedings of the International Conference on Image Processing and Computer Vision*, 2006.

[37] M. Ebrahimi and A.L. Martel. A general PDE-framework for registration of contrast enhanced images. In *Proceedings of International Conference on Medical Image Computing and Computer Assisted Intervention*, volume 12, pages 811–819, 2009.

[38] E. Ardizzone, O. Gambino, M. La Cascia, L. Lo Presti, and R. Pirrone. Multi-modal non-rigid registration of medical images based on mutual information maximization. In *Proceedings of the International Conference on Image Analysis and Processing*, pages 743–750, 2007. 7

[39] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, and D. Hawkes. Nonrigid registration using free-form deformations: application to breast mr images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999. 5, 7

[40] M. Mellor and M. Brady. Non-rigid multimodal image registration using local phase. In *LNCS: Medical Image Computing and Computer-Assisted Intervention*, pages 789–796, 2004. 6, 7, 16, 18

[41] M. Mellor and M. Brady. Phase mutual information as a similarity measure for registration. *Medical Image Analysis*, 9:330–343, 2005. 7, 16, 18, 71

[42] M. Ali and D. Clausi. Automatic registration of sar and visible band remote sensing images. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, pages 1331–1333, 2002. 7, 16

[43] F. Eugenio, F. Marques, and J. Marcello. A contour-based approach to automatic and accurate registration of multitemporal and multisensor satellite imagery. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, volume 6, pages 3390–3392, 2002. 7

[44] V. Govindu and C. Shekhar. Alignment using distributions of local geometric properties. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(10):1031–1043, 1999. 7, 16

[45] H. Li, B. Manjunath, and S. Mitra. A contour-based approach to multisensor image registration. *IEEE Transactions on Image Processing*, 4(3):320–334, 1995. 16

[46] D. Xiaolong and S. Khorram. A feature-based image registration algorithm using improved chain-code representation combined with invariant moments. *IEEE Transactions on Geoscience and Remote Sensing*, 37(5):2351–2362, 1999. 7

[47] N. Netanyahu, J. Le Moigne, and J. Masek. Georegistration of landsat data via robust matching of multiresolution features. *IEEE Transactions on Geoscience and Remote Sensing*, 42(7):1586–1600, 2004.

[48] A. Wong, W. Bishop, and J. Orchard. Efficient multi-modal least-squares alignment of medical images using quasi-orientation maps. In *Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition*, June 2006. 7, 16

[49] A. Wong and W. Bishop. Efficient least squares fusion of MRI and CT images using a phase congruency model. *Pattern Recognition Letters*, 29(3):173–180, 2008.

[50] J. Orchard. Globally optimal multimodal rigid registration: An analytic solution using edge information. In *Proceedings of the IEEE International Conference on Image Processing*, 2007. 6, 7, 16

[51] D. Russakoff, C. Tomasi, T. Rohlfing, and C. Maurer. Image similarity using mutual information of regions. In *LNCS: Proceedings of ECCV 2004*, volume 3023, pages 596–607, 2004. 9

[52] D. Field. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of The Optical Society of America A*, 4(12):2379–2394, 1987. 10, 12, 13

[53] S. Fischer, F. Sroubek, L. Perrinet, R. Redondo, and G. Cristoba. Self-invertible 2D log-gabor wavelets. *International Journal of Computer Vision*, 75(2):231–246, 2007. 10

[54] P. Kovesi. Image features from phase congruency. *Videre: A Journal of Computer Vision Research*, 1(3):2–26, 1999. 11, 16

[55] M. Morrone and R. Owens. Feature detection from local energy. *Pattern Recognition Letters*, 6:303–313, 1987. 16, 19, 25

[56] J. Fauqueur, N. Kingsbury, and R. Anderson. Multiscale keypoint detection using the dual-tree complex wavelet transform. In *Proceedings of the IEEE International Conference on Image Processing*, 2006. 18

[57] P. Kovesi. Phase congruency detects corners and edges. In *Proceedings of the Australian Pattern Recognition Society Conference*, pages 309–318, 2003. 17, 19

[58] I. Selesnick, W. Ivan, R. Baraniuk, and N. Kingsbury. The dual-tree complex wavelet transform. *IEEE Signal Processing Magazine*, 22(6):123–151, 2005. 18

[59] J. Liu, B. Vemuri, and J. Marroquin. Local frequency representations for robust multimodal image registration. *IEEE Transactions on Medical Imaging*, 21(5):462–469, 2002. 18

[60] M. Hemmendorff, M. Andersson, T. Kronander, and H. Knutsson. Phase-based multidimensional volume registration. *IEEE Transactions on Medical Imaging*, 21(12):1536–1543, 2002. 16, 18

[61] E. Adelson and J. Bergen. Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2:284–299, 1985. 16

[62] J. Bergen and E. Adelson. Early vision and texture perception. *Nature*, 333:363–364, 1988.

[63] M. Morrone and D. Burr. Feature detection in human vision: A phase-dependent energy model. *Proceedings of Royal Society of London B*, 235:221–245, 1988. 16, 19

[64] S. Venkatesh and R. Owens. An energy feature detection scheme. In *Proceedings of The International Conference on Image Processing*. 17

[65] P. Kovesi. Phase congruency: A low-level image invariant. *Psychological Research*, 64(2):136–148, 2000. 17, 25, 26, 28, 30, 34

[66] A. Wong and J. Orchard. Robust multimodal registration using local phase coherence representations. *Journal of Signal Processing Systems*, 54(1):89–100, 2009. 18

[67] Z. Wang and E. Simoncelli. Local phase coherence and the perception of blur. *Advances in Neural Information Processing Systems*, 16, May 2004. 19

[68] J. Bushberg, J. Seibert, E. Leidholdt Jr., and J. Boone. *Essential physics of medical imaging*. Williams & Wilkins, 2002. 23

[69] A. Wade and F. Fitzke. A fast, robust pattern recognition asystem for low light level image registration and its application to retinal imaging. *Optics Express*, 3(5):190–197, 1998. 23

[70] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004. 27, 32, 56

[71] D. Field and N. Brady. Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes. *Vision Research*, 37(23):3367–3383, 1997. 29

[72] T. Lindeberg. Scale-space: A framework for handling image structures at multiple scales. In *Proceedings of CERN School of Computing, Egmond aan Zee, The Netherlands*. 38

[73] A. Witkin. Scale-space filtering. In *7th International Joint Conference on Artificial Intelligence*, pages 1019–1022, 1983. 38, 39

[74] J. Koenderink and A. Van Doorn. The structure of images. *Biological Cybernetics*, pages 363–370, 1984. 38, 39, 42

[75] A. Hummel. Representations based on zero-crossings in scale space. In *Proceedings of the IEEE Computer Vision and Pattern Recognition Conference*, pages 204–209, June 1986. 39

[76] J. Babaud, A. Witkin, M. Baudin, and R. Duda. Uniqueness of the gaussian kernel for scale-space filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8, 1986. 39

[77] T. Lindeberg. Scale-space for discrete signals. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-12(3):234–254, 1990. 39

[78] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990. 40, 42, 48

[79] R. Whitaker and S. Pizer. A multi-scale approach to nonuniform diffusion. *Computer Vision, Graphics, and Image Processing*, 57(1):99–9660, 1993.

[80] L. Florack, A. Salden, B. ter Haar Romeny, J. Koenderink, and M. Viergever. Nonlinear scale-space. *Image and Vision Computing*, 13(4):279–294, 1995.

[81] Joachim Weickert and Brahim Benhamouda. A semidiscrete nonlinear scale-space theory and its relation to the perona-malik paradox. *Advances in computer vision*, pages 1–10, 1997.

[82] A. Salden. Bluman and kumei's nonlinear scale-space theory. Technical Report, February 1999. 40

[83] P. Huber. *Robust Statistics*. Wiley-Interscience, New Jersey, 1981. 64

[84] F. Hampel, E. Ronchetti, P. Rousseeuw, and W. Stahel. *Robust statistics: The Approach Based on Influence Functions*. John Wiley and Sons, New York, 1986. 65, 66

[85] M. Black and A. Rangarajan. The outlier process: Unifying line processes and robust statistics. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 15–22, 1994.

[86] M. Black and A. Rangarajan. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *International Journal of Computer Vision*, 19(1):57–91, 1996. 65

[87] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling. *Numerical Recipes in FORTRAN: The Art of Scientific Computing*. Cambridge University Press, Cambridge, 1992. 66, 68

[88] T. Hebert and R. Leahy. A generalized em algorithm for 3-D bayesian reconstruction from poisson data using gibbs priors. *IEEE Transactions on Medical Imaging*, 8:194–202, 1990. 66, 68

[89] S. Geman and D. McClure. Statistical methods for tomographic image reconstruction. *Bulletin of the International Statistical Institute*, LII-4:5–21, 1987. 67, 68

[90] K. Johnson and J. Becker. The whole brain atlas. World Wide Web Document, February 2009. `http://www.med.harvard.edu/AANLIB/home.html`. 105

[91] A. Wong, D.A. Clausi, and P. Fieguth. Adaptive monte carlo retinex method for illumination and reflectance separation and color image enhancement. In *Proceedings of the Canadian Conference on Computer and Robot Vision (CRV)*, 2009. 113

[92] A. Wong, P. Yu, W. Zhang, and D.A. Clausi. Icesynth ii: Synthesis of sar sea-ice imagery using region-based posterior sampling. *IEEE Geosciences and Remote Sensing Letters*, 7(2):348–351, 2010. 113

[93] A. Mishra, A. Wong, D.A. Clausi, and P. Fieguth. Quasi-random nonlinear scale space. *Pattern Recognition Letters*, 2010. 113

[94] A. Wong and P. Fieguth. A new bayesian source separation approach to blind decorrelation of sar data. In *Proceedings of the IEEE International Geoscience and Remote Sensing Syposium*, 2010. 113

[95] A. Wong, A. Mishra, D.A. Clausi, and P. Fieguth. Quasi-random scale space approach to robust keypoint extraction in high-noise environments. In *Proceedings of the Canadian Conference on Computer and Robot Vision (CRV)*, 2010. 114

[96] W. Zhang, A. Wong, and D.A. Clausi. JEDI: Joint enhancement and despeckling of images. In *Proceedings of the Canadian Conference on Computer and Robot Vision (CRV)*, 2009. 114

[97] A. Wong, A. Mishra, P. Fieguth, and D.A. Clausi. An adaptive monte carlo approach to nonlinear image denoising. In *Proceedings of the International Conference on Pattern Recognition (ICPR)*, 2008. 114

[98] A. Wong and J. Scharcanski. Phase-adaptive superresolution of mammographic images using complex wavelets. *IEEE Transactions on Image Processing*, 18(15):1140–1146, 2010. 114

[99] A. Wong. Adaptive bilateral filtering of image signals using local phase characteristics. *Signal Processing*, 88(6):1615–1619, 2008. 114

[100] A. Wong, D.A. Clausi, and P. Fieguth. Phase-adaptive image signal fusion using complex-valued wavelets. In *Proceedings of the International Conference on Pattern Recognition (ICPR)*, 2008. 114

[101] A. Wong. Illumination invariant active contour-based segmentation using complex-valued wavelets. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, 2008. 115